А. Г. ЦЫПКИН

СПРАВОЧНИК ПО МАТЕМАТИКЕ

ДЛЯ СРЕДНИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

А. Г. ЦЫПКИН

СПРАВОЧНИК ПО МАТЕМАТИКЕ

ДЛЯ СРЕДНИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

ИЗДАНИЕ ЧЕТВЕРТОЕ,
ПЕРЕРАБОТАННОЕ И ДОПОЛНЕННОЕ

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1988

ББК 22.1
Ц97
УДК 51(03)

Цыпкин А. Г. Справочник по математике для средних учебных заведений. – 4-е изд., испр. и доп. – М.: Наука. Гл. ред. физ.-мат. лит., 1988. – 432 с.

Справочник предназначен для учащихся средних школ и средних специальных учебных заведений. Он содержит все необходимые определения, формулы, теоремы и методы решения задач. В него включены, помимо классических разделов элементарной математики, такие разделы, как элементы теории множеств, комплексные числа, основы математического анализа и векторной алгебры, метод координат и т. д. Материал, излагаемый в справочнике, в основном носит теоретический характер. Новое издание справочника (3-е выходило в 1983 г.) дополнено разделами, которые внедряются в школьное обучение. К ним, в частности, относится теория вероятностей.

Ц $\frac{1702010000\text{-}001}{053(02)\text{-}88}$ 61-88

ISBN 5-02-013792-8

© Издательство «Наука». Главная редакция физико-математической литературы, 1979; с изменениями, 1983; с изменениями, 1988

ОГЛАВЛЕНИЕ

От автора 9

Глава 1. **Элементы теории множеств** 10

§ 1. Множества и операции над множествами 10
1.1. Множества и подмножества (10). 1.2. Операции над множествами (11). 1.3. Упорядоченные множества (13).

§ 2 Соответствие между множествами и отображение множеств 14
2.1. Соответствие и отображение (14). 2.2. Взаимно однозначное отображение (15). 2.3. Эквивалентность множеств (15). 2.4. Классификация множеств (16).

§ 3. Множества с бинарными операциями 18
3.1. Бинарные операции в множествах (18). 3.2. Изоморфизм множеств (10). 3.3. Группы (20). 3.4. Кольца (21). 3.5. Поля (21).

Глава 2. **Действительные числа** 22

§ 1. Натуральные числа 23
1.1. Множество натуральных чисел (23). 1.2. Аксиоматическое построение множества натуральных чисел (25). 1.3. Простые числа. Основная теорема арифметики (26). 1.4. Некоторые признаки делимости натуральных чисел (27). 1.5. Наименьшее общее кратное. Наибольший общий делитель. Алгоритм Евклида (27).

§ 2. Целые числа 30
2.1. Множество целых чисел (30). 2.2. Арифметические операции с целыми числами (30). 2.3. Множество целых чисел как множество упорядоченных пар натуральных чисел (32).

§ 3. Рациональные числа 35
3.1. Рациональные дроби. (35). 3.2. Рациональные числа (38) 3.3. Множество рациональных чисел как множество упорядоченных пар целых чисел (39).

§ 4. Десятичные дроби 40
4.1. Десятичная позиционная система счисления (40). 4.2. Понятие десятичной дроби (41) 4.3. Арифметические действия с конечными десятичными дробями (43). 4.4. Обращение конечной десятичной дроби в рациональную дробь (44). 4.5. Обращение бесконечной периодической дроби в рациональную дробь (45). 4.6. Десятичные представления рациональных чисел (46). 4.7. Непрерывные дроби (48).

§ 5. Действительные числа 51
5.1. Множество действительных чисел как расширение множества рациональных чисел (51). 5.2. Аксиоматическое построение множества действительных чисел (52). 5.3. Представление действительных чисел десятичными дробями (54). 5.4. Геометрическое изображение множества действительных чисел (56). 5.5. Степени и корни (57). 5.6. Логарифмы (58). 5.7. Пропорции (58). 5.8. Десятичные представления иррациональных чисел (59). 5.9. Некоторые способы доказательства иррациональности чисел (59). 5.10. Алгебраические и трансцендентные числа (62).

§ 6. Приближенные вычисления 63
6.1. Приближенное значение числа и погрешности (63). 6.2. Десятичная запись приближенных значений числа (64). 6.3. Округление чисел (66). 6.4. Метод касательных (метод Ньютона) (66). 6.5. Алгоритм извлечения квадратного корня (68).

Глава 3. **Комплексные числа** 70

§ 1. Множество комплексных чисел 71
1.1. Аксиоматическое построение множества комплексных чисел (71). 1.2. Множество комплексных чисел как множество упорядоченных пар действительных чисел (73).

§ 2. Геометрическое изображение и тригонометрическая форма записи комплексных чисел 75
2.1. Геометрическое изображение комплексного числа (75). 2.2. Геометрическое изображение суммы и разности комплексных чисел (76). 2.3. Тригонометрическая форма записи комплексного числа (77).

§ 3. Степень комплексного числа 78
3.1. Натуральная степень комплексного числа (78). 3.2. Корень n-й степени из комплексного числа (78).

Глава 4. **Алгебра** 80

§ 1. Многочлены одного переменного 80
1.1. Понятие многочлена. Арифметические операции над многочленами (80). 1.2. Делители многочлена (82). 1.3. Деление многочленов. Схема Горнера. Теорема Безу (83). 1.4. Алгоритм Евклида (85). 1.5. Корни многочлена (86). 1.6. Формулы сокращенного умножения (86). 1.7. Формулы Виета (87). 1.8. Основная теорема алгебры (87) 1.9. Разложение многочлена на множители (88). 1.10. Некоторые следствия основной теоремы алгебры (88).

§ 2. Многочлены нескольких переменных 89
2.1. Одночлены и многочлены нескольких переменных (89). 2.2. Лексикографическое расположение членов многочлена (90).

§ 3. Рациональные алгебраические дроби 91
3.1. Множество рациональных алгебраических дробей (91). 3.2. Правильные алгебраические дроби (94). 3.3. Простейшие дроби (94).

§ 4. Иррациональные алгебраические выражения 96

§ 5. Уравнения. Алгебраические уравнения 98
5.1. Основные определения (98). 5.2. Линейное уравнение (100). 5.3. Квадратное уравнение (100). 5.4. Кубичное уравнение (102). 5.5. Уравнение четвертой степени (102). 5.6. Некоторые уравнения четвертой степени специального вида (103). 5.7. Двучленные уравнения (105). 5.8. Решение алгебраического уравнения с целыми коэффициентами (106). 5.9. Рациональные алгебраические уравнения (107). 5.10. Иррациональные уравнения (107). 5.11. Уравнения, содержащие неизвестное под знаком абсолютной величины (110). 5.12. Решение уравнений в множестве комплексных чисел (111). 5.13. Диофантовы уравнения (112).

§ 6. Трансцендентные уравнения 113
6.1. Показательные уравнения (113). 6.2. Логарифмические уравнения (114).

§ 7. Системы уравнений 116
7.1. Основные определения (116). 7.2. Системы линейных уравнений (117). 7.3. Метод последовательного исключения неизвестных (метод Гаусса) (117). 7.4. Некоторые способы решениясистем нелинейных алгебраических уравнений (122).

§ 8. Исследование систем линейных уравнений с помощью определителей 124
8.1. Матрицы и операции над ними (124). 8.2. Определители (127). 8.3. Ранг матрицы (129). 8.4. Исследование систем линейных уравнений с помощью определителей. Формулы Крамера (131).

§ 9. Неравенства 137
9.1. Определения и основные свойства неравенств (137). 9.2. Некоторые важные неравенства (138).

§ 10. Решение неравенств и систем неравенств 140
10.1. Основные определения (140). 10.2. Линейные неравенства и системы неравенств с одним неизвестным (140). 10.3. Квадратные неравенства (141). 10.4. Метод интерва-

лов (142). 10.5. Решение иррациональных неравенств (143). 10.6. Показательные неравенства (144). 10.7. Логарифмические неравенства (145). 10.8. Геометрическое изображение множества решений неравенства с двумя неизвестными (147).

§ 11. Приемы доказательства справедливости неравенств . . 151
11.1. Доказательство с помощью цепочки эквивалентных неравенств (151). 11.2. Доказательство с использованием свойств функций, входящих в неравенства (152). 11.3. Некоторые специальные приемы доказательства (153). 11.4. Некоторые способы проверки справедливости числовых неравенств (154).

Глава 5. Геометрия 157

§ 1. Луч. Отрезок 157
1.1. Луч (157). 1.2. Отрезок (159).

§ 2. Углы на плоскости 159
2.1. Понятие угла (159). 2.2. Градусная мера измерения углов (160). 2.3. Радианная мера измерения углов (160). 2.4. Классификация углов (160). 2.5. Угол между направлениями (161).

§ 3. Параллельность и перпендикулярность на плоскости 161
3.1. Параллельность на плоскости (161). 3.2. Перпендикулярность на плоскости (162). 3.3. Расстояние от точки до прямой (162).

§ 4. Параллельность и перпендикулярность в пространстве 163
4.1. Параллельность прямой и плоскости (163). 4.2. Параллельность плоскостей (163). 4.3. Перпендикулярность прямой и плоскости (163). 4.4. Расстояние от точки до плоскости (164). 4.5. Перпендикулярность плоскостей (164). 4.6. Наклонная (164). 4.7. Скрещивающиеся прямые (164).

§ 5. Проектирование на плоскость 165
5.1. Параллельное проектирование (165). 5.2. Ортогональное проектирование (165).

§ 6. Углы в пространстве 165
6.1. Угол между наклонной и плоскостью (165). 6.2. Двугранный угол (166). 6.3. Угол между двумя плоскостями (166).

§ 7. Ломаная. Многоугольник 166

§ 8. Треугольники 168
8.1. Основные свойства (168). 8.2. Медианы треугольника (169). 8.3. Высоты треугольника (170). 8.4. Биссектрисы треугольника (170). 8.5. Средняя линия треугольника (171). 8.6. Равнобедренный треугольник (171). 8.7. Равносторонний треугольник (171). 8.8. Прямоугольный треугольник (172).

§ 9. Четырехугольники 172
9.1. Параллелограмм (172). 9.2. Ромб (173). 9.3. Прямоугольник (174). 9.4. Квадрат (174). 9.5. Трапеция (174).

§ 10. Подобные многоугольники 175
10.1. Признак подобия многоугольников (175). 10.2. Признаки подобия треугольников (175).

§ 11. Окружность и круг 176
11.1. Окружность и круг (176). 11.2. Касательная и секущая (177). 11.3. Взаимное расположение двух окружностей (177). 11.4. Центральные углы и дуги окружности. (178). 11.5. Дуги и хорды окружности (179). 11.6. Углы в окружности (179). 11.7. Длины и площади в окружности и круге (180).

§ 12. Многоугольники и окружности 181
12.1. Вписанные и описанные многоугольники (181). 12.2. Вписанные треугольники (182). 12.3. Описанные треугольники (182). 12.4. Вневписанная окружность (182). 12.5. Соотношения между сторонами треугольников и радиусами вписанной и описанной окружностей (183). 12.6. Вписанные четырехугольники (183). 12.7. Описанные четырехугольники (183).

§ 13. Геометрические построения 184
13.1. Построение прямых, параллельных и перпендикулярных данной прямой (184). 13.2. Построение углов (185). 13.3. Построение отрезков (186). 13.4. Построение окружностей и дуг окружностей (189). 13.5. Построение касательных к окружностям (191). 13.6. Построение окружности, описанной около многоугольника, и многоугольника, вписанного в окруж-

ность (192). 13.7. Построение окружности, вписанной в многоугольник, и многоугольника, описанного около окружности (193). 13.8. Построение треугольников (194).

§ 14. Многогранный угол 197
§ 15. Многогранная поверхность. Многогранник 198
§ 16. Призма 199
§ 17. Параллелепипед. Куб 200
§ 18. Пирамида. Усеченная пирамида 201
§ 19. Правильные многогранники 203
§ 20. Цилиндр 205
§ 21. Конус. Усеченный конус 207
§ 22. Сфера. Шар 209
§ 23. Части шара 211
23.1. Шаровой сегмент (211). 23.2. Шаровой сектор (211). 23.3. Шаровой слой (211). 23.4. Шаровой пояс (212). 23.5. Телесный угол (212).
§ 24. Преобразования плоскости и пространства 213
24.1. Геометрические фигуры. Отображение фигуры в фигуру и отображение фигуры на фигуру (213). 24.2. Преобразование плоскости и пространства (213). 24.3. Изометрия пространства и плоскости. Равенство фигур (214). 24.4. Поворот плоскости вокруг точки (215). 24.5. Центральная симметрия и центрально симметричные фигуры (215). 24.6. Осевая симметрия плоскости (215). 24.7. Осевая симметрия пространства (216). 24.8. Симметрия относительно плоскости (217). 24.9. Гомотетия плоскости (218). 24.10. Гомотетия пространства (219). 24.11. Преобразование подобия плоскости (219). 24.12. Подобные фигуры (220).
§ 25. Система аксиом и неопределяемых понятий геометрии (по Гильберту) 220

Глава 6. **Тригонометрия** 224
§ 1. Тригонометрические функции 224
1.1. Обобщение понятия угла (224). 1.2. Тригонометрические функции (225). 1.3. Квадранты единичной окружности. Знаки значений тригонометрических функций (227). 1.4. Тригонометрические функции числового аргумента (227). 1.5. Обратные тригонометрические функции (232). 1.6. Значения тригонометрических функций некоторых углов (234).
§ 2. Тригонометрические формулы 237
2.1. Формулы приведения (237). 2.2. Связь между тригонометрическими функциями одного аргумента (237). 2.3. Тригонометрические функции суммы и разности углов (239). 2.4. Тригонометрические функции двойных, тройных и половинных углов (239). 2.5. Преобразование суммы (разности) тригонометрических функций в произведение (239). 2.6. Преобразование произведения тригонометрических функций в сумму (240). 2.7. Простейшие соотношения между обратными тригонометрическими функциями (240).
§ 3. Решение тригонометрических уравнений и неравенств 241
3.1. Простейшие тригонометрические уравнения (241). 3.2. Примеры более сложных тригонометрических уравнений (243). 3.3. Решение простейших тригонометрических неравенств (246). 3.4. Примеры решения уравнений и неравенств, содержащих обратные тригонометрические функции (246).
§ 4. Соотношения между элементами треугольника 249
4.1. Основные формулы (249). 4.2. Вычисление элементов треугольника (250).

Глава 7. **Метод координат** 254
§ 1. Системы координат 254
1.1. Координатная ось (254). 1.2. Прямоугольная декартова система координат на плоскости (255). 1.3. Полярная система координат. Связь между прямоугольными и полярными координатами (257). 1.4. Прямоугольная декартова система координат в пространстве (258).

§ 2. Линии на плоскости и поверхности в пространстве . . 259
2.1. Линии первого и второго порядка (259). 2.2. Прямая (260). 2.3. Окружность (262). 2.4. Эллипс (262). 2.5. Гипербола (264). 2.6. Парабола (266). 2.7. Плоскость. Прямые в пространстве (267).
§ 3. Векторы 269
3.1. Векторы. Основные понятия (269). 3.2. Угол между векторами. Скалярное произведение векторов (274). 3.3. Координаты вектора на плоскости (274). 3.4. Координаты вектора в пространстве (275). 3.5. Векторное произведение (276). 3.6. Смешанное (скалярно-векторное) произведение векторов (277).

Глава 8. Теория пределов 279
§ 1. Числовые последовательности 279
1.1. Понятие числовой последовательности (279). 1.2. Некоторые способы задания последовательностей (280). 1.3. Геометрическое изображение членов последовательности (281). 1.4. Ограниченные последовательности (281). 1.5. Монотонные последовательности (282).
§ 2. Предел последовательности 283
2.1. Понятие предела последовательности (283). 2.2. Необходимое условие сходимости последовательности (285). 2.3. Теоремы о пределах последовательностей (286). 2.4. Достаточное условие сходимости последовательности (теорема Вейерштрасса) (288). 2.5. Необходимое и достаточное условие сходимости последовательности (292).
§ 3. Числовые ряды и бесконечные произведения 293
3.1. Понятие числового ряда (293). 3.2. Положительные числовые ряды (295). 3.3. Признаки сходимости числовых рядов (297). 3.4. Некоторые конечные суммы и числовые ряды (298). 3.5. Понятие бесконечного произведения (299). 3.6. Связь бесконечных произведений с рядами (300).
§ 4. Прогрессии 300
4.1. Арифметическая прогрессия (300). 4.2. Геометрическая прогрессия (301).
§ 5. Числовые функции 302
5.1. Понятие числовой функции (302). 5.2. Способы задания функций (303). 5.3. Сумма, произведение, разность и частное двух функций (304). 5.4. Сложная функция (суперпозиция функций) (304). 5.5. Четные и нечетные функции (305). 5.6. Периодические функции (306). 5.7. Ограниченные функции (307). 5.8. Монотонные функции (307). 5.9. Понятие обратной функции (309). 5.10. Понятие числовой функции нескольких независимых переменных (310).
§ 6. Предел функции 311
6.1. Понятие предела функции (311). 6.2. Односторонние пределы (314). 6.3. Теоремы о пределах функций (315). 6.4. Пределы некоторых функций (317).
§ 7. Бесконечно малые и бесконечно большие величины . . 319
7.1. Понятие бесконечно малой и бесконечно большой величины (319). 7.2. Сравнение бесконечно малых и бесконечно больших величин (321).
§ 8. Непрерывность (и разрывы) функций 322
8.1. Понятие непрерывности функций (322). 8.2. Основные теоремы о непрерывных функциях. Непрерывность элементарных функций (323). 8.3. Односторонняя непрерывность. Классификация разрывов функций (325). 8.4. Предел и непрерывность функций двух независимых переменных (325).

Глава 9. Начала дифференциального и интегрального исчислений 327
§ 1. Производная 327
1.1. Понятие производной (327). 1.2. Геометрическая интерпретация производной (329). 1.3. Кинематический смысл производной (330). 1.4. Свойства производных и основные теоремы дифференциального исчисления (330). 1.5. Вычисление производных элементарных функций (331). 1.6. Производные вы-

шего порядка (334). 1.7. Производные функции, заданной в параметрической форме (336). 1.8. Понятие частной производной (337).

§ 2. Дифференциал 339
2.1. Понятие дифференциала (339). 2.2. Геометрическая интерпретация дифференциала (340). 2.3. Инвариантность дифференциала (340).

§ 3. Исследование функций. Построение графика функций 341
3.1. Условия постоянства и монотонности функций (341). 3.2. Максимумы и минимумы функции (342). 3.3. Наибольшее и наименьшее значения функции (344). 3.4. Направление вогнутости кривой (345). 3.5. Асимптоты графика функции (346). 3.6. Построение графика функции (347). 3.7 Экстремумы функции двух независимых переменных (350).

§ 4. Первообразная. Неопределенный интеграл 351
4.1. Понятия первообразной и неопределенного интеграла (351). 4.2. Простейшие правила интегрирования (353). 4.3. Метод интегрирования по частям (353). 4.4. Интегрирование методом замены переменной (356). 4.5. Некоторые классы интегрируемых функций (358).

§ 5. Определенный интеграл 363
5.1. Задача вычисления площади плоской фигуры (363). 5.2. Понятие определенного интеграла (363). 5.3. Определенный интеграл как функция верхнего предела. Основная формула интегрального исчисления (366). 5.4. Вычисление площадей и объемов с помощью определенного интеграла (367).

§ 6. Дифференциальные уравнения 370
6.1 Понятие дифференциального уравнения (370). 6.2. Дифференциальные уравнения первого порядка (371). 6.3. Некоторые дифференциальные уравнения первого порядка (373). 6.4. Дифференциальные уравнения порядка выше первого. Примеры дифференциальных уравнений второго порядка (377). 6.5. Линейные дифференциальные уравнения с постоянными коэффициентами (379).

Глава 10. **Комбинаторика. Элементы теории вероятностей** 384

§ 1. Метод математической индукции 384

§ 2. Комбинаторика 386
2.1. Перестановки (386). 2.2. Подстановки (388). 2.3. Размещения (389). 2.4. Сочетания (390). 2.5. Бином Ньютона (391). 2.6. Перестановки с повторениями (392). 2.7. Сочетания с повторениями (394).

§ 3. Элементы теории вероятностей 395
3.1. Испытание. Исход. Событие (395). 3.2. Операции над событиями (396). 3.3. Понятие вероятности (397). 3.4. Условная вероятность (400). 3.5. Формула полной вероятности (401). 3.6. Формула Бернулли (402).

Приложения 404

I. Элементарные функции 404

II. Системы счисления 412

Список основных обозначений 417

Предметный указатель 419

ОТ АВТОРА

Справочник предназначен для учащихся средних школ и средних специальных учебных заведений и в полном объеме содержит понятия, определения, формулы, теоремы и методы решения задач, входящие в курс математики для средних учебных заведений. Кроме того, в справочник дополнительно включен ряд разделов, необходимых как для более глубокого понимания основ математики, так и для выявления связей между различными разделами математики. К дополнительному материалу, в частности, относятся: алгоритм Евклида, комплексные числа, основная теорема алгебры, кривые второго порядка, понятия эквивалентности и изоморфизма множеств, элементы теории вероятностей, начала теории дифференциальных уравнений, основные сведения о матрицах и определителях, представление о структуре и принципах построения действительных чисел. В справочнике теория действительных чисел излагается с единой точки зрения, а именно множества целых, рациональных и действительных чисел последовательно вводятся как естественное расширение множества натуральных чисел.

Несколько расширен круг вопросов, связанных с приближенными вычислениями. Наряду с элементарными сведениями общего характера в справочнике рассмотрены аппроксимирующие свойства непрерывных дробей и изложен один из наиболее простых и известных методов приближенного вычисления — метод касательных Ньютона.

Обозначения, принятые в справочнике, за редким исключением совпадают с обозначениями, принятыми в школьных учебниках. К указанным исключениям относится обозначение интервала: в справочнике интервал обозначается круглыми скобками, как это принято в большинстве учебников и книг по математике.

В справочнике принята следующая рубрикация: глава, параграф, пункт; названия всех пунктов указаны в оглавлении. Сплошная нумерация рисунков ведется в пределах одной главы. При ссылке в тексте на материал той же самой главы указывается только номер параграфа и пункта; при ссылке на материал другой главы дополнительно указывается номер главы. В конце справочника дан предметный указатель.

Четвертое издание справочника по структуре изложения материала незначительно отличается от третьего издания, вышедшего в 1983 г. Некоторые разделы существенно переработаны. Добавлены элементы комбинаторики и теории вероятностей.

Автор надеется, что справочник в предлагаемом виде окажется более удобным для читателей и может быть использован не только для получения кратких сведений об основных понятиях, формулах и теоремах, но также и для повторения курса математики средней школы.

Глава 1

ЭЛЕМЕНТЫ ТЕОРИИ МНОЖЕСТВ

Теория множеств — это раздел математики, изучающий общие свойства множеств (преимущественно бесконечных). Выделение теории множеств в самостоятельный раздел математики произошло сравнительно недавно — на рубеже XIX и XX веков. Теория множеств оказала очень большое влияние на развитие современной математики — она явилась фундаментом ряда новых разделов, позволила по-новому взглянуть на классические разделы и глубже понять сам предмет математики.

§ 1. Множества и операции над множествами

1.1. Множества и подмножества. В математике некоторые понятия являются первичными, неопределяемыми. К ним относятся понятия натурального числа, точки, прямой и т. д.

Одним из таких неопределяемых понятий является понятие «множество». Этому понятию нельзя дать формального определения, которое не сводилось бы просто к замене слова «множество» его синонимами «совокупность», «набор элементов» и т. п. *Множества* можно составлять на основе самых различных признаков из самых разнообразных объектов (которые в дальнейшем будем называть *элементами* множества). Множество можно задать, указывая, например, некоторое свойство объектов, образующих множество, правило построения элементов множества и т. д. Можно говорить не только о множествах, элементами которых являются материальные объекты, но и о множествах, элементы которых — некоторые абстрактные понятия (числа, геометрические фигуры, символы и т. п.).

Сразу оговорим, что понятие «множество» не следует понимать буквально и толковать его как совокупность, содержащую «много» элементов. Под множеством также понимается совокупность объектов, которая может состоять, например, из одного, двух и т. д. элементов. Более того, оказывается удобным считать множеством даже *пустое* множество, т. е. множество, не содержащее ни одного элемента. Рассматривать пустое множество необходимо хотя бы только потому, что, когда мы определяем тем или иным способом множество, мы можем и не знать заранее, содержит ли оно хотя бы один элемент.

Множества чаще всего обозначаются прописными буквами латинского алфавита: A, B, ..., X, а их элементы — строчными буквами: a, b, ..., x. Пустое множество обозначается специальным символом $\varnothing$.

Если множество A состоит из n элементов a_1, a_2, ..., a_n, то пишут

$$A = \{a_1;\ a_2;\ \ldots;\ a_n\}.$$

Говорят, что «элемент a принадлежит множеству A» — и записывают: $a \in A$ или $A \ni a$ (A содержит a); запись $a \notin A$ или $A \not\ni a$ означает, что элемент a не принадлежит множеству A (A не содержит a).

Множество B называют *подмножеством* множества A, если все элементы множества B принадлежат множеству A, и пишут

$$B \subset A.$$

Например, пусть A — множество рациональных чисел, B — множество натуральных чисел. В этом случае $B \subset A$.

Любое множество A имеет в качестве своих подмножеств пустое множество и само множество A.

Если для двух множеств A и B одновременно справедливы утверждения $A \subset B$ и $B \subset A$, то множества A и B состоят из одних и тех же элементов. Такие множества A и B называют *равными* (или *совпадающими*) и пишут

$$A = B.$$

Непустое подмножество B множества A называется *собственным*, если B не совпадает с A.

1.2. Операции над множествами. *Объединением* двух множеств называется множество, состоящее из всех элементов, принадлежащих

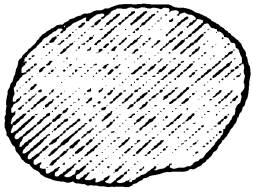


Рис. 1.1

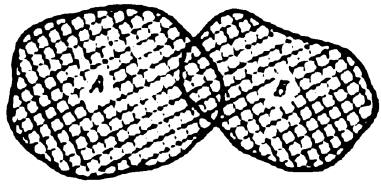


Рис. 1.2

хотя бы одному из этих множеств. *Объединение* множеств A и B обозначается $A \cup B$. Для наглядности изобразим множество в виде плоской геометрической фигуры (на рис. 1.1. элементы множества — точки заштрихованной части плоскости). На рис. 1.2 множество $A \cup B$ выделено двойной штриховкой.

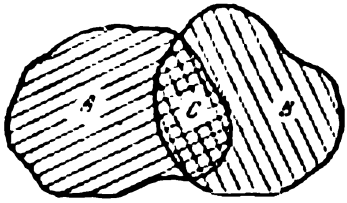


Рис. 1.3

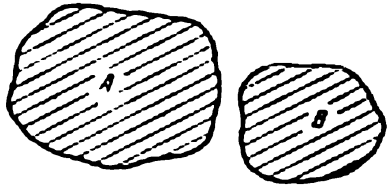


Рис. 1.4

Пересечением двух множеств называется множество, состоящее из всех элементов, принадлежащих одновременно обоим множествам. Пересечение множеств A и B обозначается $A \cap B$. На рис. 1.3 множество $C = A \cap B$ выделено двойной штриховкой.

Множества A и B могут быть такими, что их пересечение будет пустым множеством: $A \cap B = \varnothing$ (рис. 1.4).

Свойства операций объединения и пересечения:

1) коммутативность:

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A;$$

2) ассоциативность:

$$(A \cup B) \cup C = A \cup (B \cup C), \qquad (A \cap B) \cap C = A \cap (B \cap C);$$

3) **дистрибутивность**:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C),$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

Разностью множеств A и B называется множество, состоящее из всех элементов A, не принадлежащих B. При этом, вообще говоря, не предполагается, что $B \subset A$. Разность множеств A и B обозначается $A \setminus B$ (или $A - B$).

Если B — подмножество A, то разность $A \setminus B$ также называют *дополнением* множества B до множества A. Если $R = A \setminus B$, то $R \cap B = \varnothing$ и $R \cup B = A$. На рис. 1.5 дополнение $A \setminus B$ выделено двойной штриховкой.

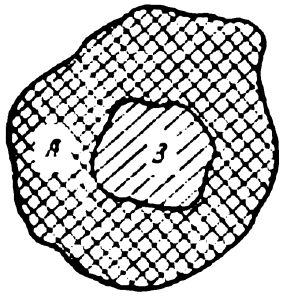


Рис. 1.5

Поясним смысл этих определений следующим примером. Пусть A — множество всех натуральных чисел, кратных двум, а B — множество всех натуральных чисел, кратных трем. Объединением этих множеств будет множество всех натуральных чисел, кратных числу 2 или числу 3. Их пересечением будет множество всех натуральных чисел, кратных как числу 2, так и числу 3, т. е. кратных числу 6.

Пусть теперь A — множество всех натуральных чисел, кратных числу 2, а B — множество всех натуральных чисел, кратных числу 6 (т. е. кратных как числу 2, так и числу 3). Множество B является подмножеством множества A. Дополнением множества B до множества A будет множество всех натуральных чисел, кратных числу 2, но не кратных числу 3.

Понятия объединения и пересечения, данные для случая двух множеств могут быть распространены и на случай любого числа множеств.

Объединением конечного числа множеств A_i ($i = 1, 2, \ldots, n$) называют множество B, состоящее из элементов, каждый из которых принадлежит хотя бы одному из множеств A_i, и пишут

$$B = \bigcup_{i=1}^{n} A_i.$$

Пересечением конечного числа множеств A_i ($i = 1, 2, \ldots, n$) называют множество C, состоящее из всех элементов, принадлежащих одновременно всем множествам A_i, и пишут

$$C = \bigcap_{i=1}^{n} A_i.$$

Если A_i — бесконечная совокупность множеств ($i = 1, 2, \ldots, n, \ldots$), то объединение и пересечение этих множеств обозначают с помощью символов

$$\bigcup_i A_i \quad \text{и} \quad \bigcap_i A_i$$

соответственно.

Пусть A_i — конечная ($i = 1, 2, \ldots, n$) или бесконечная ($i = 1, 2, \ldots, n, \ldots$) совокупность подмножеств множества A. Связь между подмножествами A_i и множеством A устанавливается следующими равенствами, называемыми

соотношениями *двойственности*

$$A \setminus \bigcap_i A_i = \bigcup_i (A \setminus A_i), \quad A \setminus \bigcup_i A_i = \bigcap_i (A \setminus A_i).$$

Пусть A и B — два произвольных множества. Пару (a, b) элементов $a \in B$ и $b \in B$, взятых в данном порядке, называют *упорядоченной парой*. Считают, что пары (a_1, b_1) и (a_2, b_2) равны тогда и только тогда, когда $a_1 = a_2$ и $b_1 = b_2$. *Декартовым произведением* двух множеств A и B называется множество всех упорядоченных пар (a, b). Декартово произведение множеств A и B обычно обозначается $A \times B$.

1.3. Упорядоченные множества. Вводя операции над множествами, мы не учитывали, что сами множества могут иметь свою внутреннюю структуру, т. е. мы считали, что все элементы множества равноправны. Однако в математике такие «чистые» множества представляют мало интереса и гораздо чаще изучаются множества, между элементами которых существуют те или иные отношения. Одним из важнейших отношений между элементами множества является отношение порядка.

Отношение порядка есть не что иное, как правило, устанавливающее порядок «следования» элементов множества.

Пусть A — некоторое множество. Множество A называется *упорядоченным множеством*, если для любых двух его элементов a, b установлено одно из *отношений порядка*:

либо $a \leqslant b$ (a не превосходит b),
либо $b \leqslant a$ (b не превосходит a),

обладающих следующими свойствами:

1) рефлексивность: любой элемент не превосходит самого себя;

2) антисимметричность: если a не превосходит b, а b не превосходит a, то элементы a и b совпадают;

3) транзитивность: если a не превосходит b, а b не превосходит c, то a не превосходит c.

Пустое множество условились считать упорядоченным. В сформулированном выше определении упорядоченного множества, элементами которого могут быть объекты любой природы, знак $\leqslant$ читается «не превосходит». Привычное чтение и смысл («меньше или равно») этот знак приобретает в случае, когда элементы множества A — числа.

Два множества, составленные из одних и тех же элементов, но с разными отношениями порядка, считаются различными упорядоченными множествами.

Одно и то же множество можно упорядочить различными способами, получая тем самым различные упорядоченные множества. Например, рассмотрим множество, элементами которого являются различные выпуклые многоугольники: треугольник, четырехугольник, пятиугольник, шестиугольник и т. д. Один способ образования упорядоченного множества из данного неупорядоченного множества может, например, состоять в том, что в качестве первого элемента упорядоченного множества мы берем треугольник, в качестве второго — четырехугольник, третьего — пятиугольник и т. д., т. е. упорядочиваем множество в порядке возрастания числа внутренних углов многоугольников. Множество многоугольников может быть упорядочено и другим способом, например перечислением многоугольников в порядке возрастания площадей, когда в качестве первого выбирается

многоугольник, имеющий наименьшую площадь, в качестве второго — многоугольник с площадью, не превышающей площадь всех остальных, кроме уже выбранного, и т. д.

Упорядоченные (конечные или счетные) множества часто записывают, располагая их элементы в заданном порядке в круглых скобках. Например, записи

$$(1; 2; 3) \quad \text{и} \quad (2; 1; 3) \tag{1}$$

представляют различные конечные упорядоченные множества, которые можно получить из одного и того же множества $\{1; 2; 3\}$, упорядочивая его двумя различными способами. Для записи счетного упорядоченного множества в виде, аналогичном (1), необходимо указать первый элемент упорядоченного множества и указать порядок (правило) следования элементов.

§ 2. Соответствие между множествами и отображение множеств

2.1. Соответствие и отображение. Понятие соответствия относится к первичным, неопределяемым понятиям математики. Говорят, что между двумя множествами установлено *соответствие*, если определено правило, по которому для каждого элемента одного множества выбирается определенный элемент или подмножество элементов другого множества. При этом допускается, что некоторым элементам первого множества может соответствовать пустое подмножество.

На основе понятия соответствия между множествами вводится понятие отображения множеств. При этом различают отображение множества «в» множество и отображение множества «на» множество.

Соответствие, при котором каждому элементу множества A отвечает единственный элемент множества B, называется *отображением множества A в множество B*.

Соответствие, при котором каждому элементу множества A отвечает единственный элемент множества B и, кроме того, каждому элементу множества B отвечает хотя бы один элемент множества A, называется *отображением множества A на множество B*.

Отображения множеств обычно обозначают буквами $f, g, h, \ldots$ и пишут

$$A \xrightarrow{f} B, \quad A \xrightarrow{g} B, \quad A \xrightarrow{h} B$$

или

$$f\colon A \to B, \quad g\colon A \to B, \quad h\colon A \to B.$$

Если при отображении f элементу $a \in A$ соответствует элемент $b \in B$, то элемент b называют *образом* элемента a, элемент a называют *прообразом* элемента b и пишут

$$b = f(a).$$

Множество всех прообразов элемента b называют его *полным прообразом*. В случае отображения множества A «на» множество B пишут также

$$B = f(A).$$

Отображение $f\colon A \to B$ называется *инъективным*, если разные элементы множества A имеют различные образы.

Равенство $f=g$ двух отображений

$$f\colon A \longrightarrow B, \quad g\colon A \longrightarrow B$$

означает, что для любого элемента $c \in A$ образы элемента a при отображениях f и g совпадают: $f(a)=g(a)$.

2.2. Взаимно однозначное отображение. Отображение множества A на множество B, при котором разным элементам множества A соответствуют разные элементы множества B, называется *взаимно однозначным отображением* множества A на множество B. Другими словами, отображение $f\colon A \longrightarrow B$ взаимно однозначно, если оно отображает множество A на множество B и оно инъективно. Взаимно однозначное отображение называют также *биекцией*.

Если множества A и B совпадают и f — взаимно однозначное отображение, т. е. $f\colon A \longrightarrow A$, то говорят, что множество A взаимно однозначно отображается на себя.

Тождественным (или *единичным*) отображением e множества A на себя:

$$e\colon A \longrightarrow A,$$

называется отображение, переводящее любой элемент $a \in A$ в себя: $e(a)=a$.

Пусть множество A взаимно однозначно отображается на множество B ($f\colon A \longrightarrow B$), т. е. каждому элементу $a \in A$ при отображении f соответствует единственный элемент b множества B (образ a) и каждому элементу $b \in B$ соответствует единственный элемент $a \in A$ (прообраз b). Отображение, ставящее в соответствие каждому элементу $b \in B$ его прообраз $a \in A$, называют *обратным отображением* для отображения f, обозначают f^{-1} и пишут

$$B \xrightarrow{f^{-1}} A \quad \text{или} \quad f^{-1}\colon B \longrightarrow A.$$

Если f — взаимно однозначное отображение, то обратное отображение f^{-1} — также взаимно однозначное отображение; обратным для отображения f^{-1} будет исходное отображение f.

2.3. Эквивалентность множеств. Если множество A взаимно однозначно отображается на множество B, то множества A и B называются *эквивалентными*. Эквивалентность множеств записывается с помощью знака $\sim$:

$$A \sim B$$

(читается «A эквивалентно B»). Об эквивалентных множествах A и B также говорят, что между ними установлено *отношение эквивалентности*.

Отношение эквивалентности множеств обладает следующими свойствами:

1) рефлексивность:

$$A \sim A$$

(любое множество эквивалентно самому себе);

2) симметричность:

$$A \sim B \Leftrightarrow B \sim A$$

(если множество A эквивалентно множеству B, то множество B эквивалентно множеству A, и наоборот);

3) транзитивность:

$$A \sim B \text{ и } B \sim C \Rightarrow A \sim C$$

(если множество A эквивалентно множеству B, а множество B эквивалентно множеству C, то и множество A эквивалентно множеству C).

2.4. Классификация множеств. Множества делятся на конечные и бесконечные.

Множество, эквивалентное отрезку натурального ряда*), называется *конечным множеством* (пустое множество также считается конечным множеством). Иными словами, конечное множество (если оно не пусто) — это множество, элементы которого можно «пересчитать» за конечное число шагов. Если при этом каждому элементу конечного множества A присвоить какой-нибудь номер от 1 до некоторого натурального числа n так, чтобы все числа от 1 до n были использованы и различные элементы множества получили различные номера, то число n укажет число элементов данного множества A.

Число элементов конечного множества A называется *мощностью* конечного множества.

Для конечных множеств справедлива следующая теорема, которая называется основной теоремой о конечных множествах:

Любое конечное множество не эквивалентно никакому его собственному подмножеству.

Из этой теоремы, в частности, также следует, что всякое непустое конечное множество эквивалентно одному и только одному отрезку натурального ряда.

Множество, не являющееся конечным, называется *бесконечным*.

В качестве «эталона» для сравнения бесконечных множеств можно выбрать простейшее бесконечное множество — множество всех натуральных чисел N. Если проводить сравнение всевозможных бесконечных множеств с множеством всех натуральных чисел N, то оказывается, что все бесконечные множества разбиваются на два класса: на класс множеств, эквивалентных множеству всех натуральных чисел (такие множества называют *счетными*), и на класс множеств, не эквивалентных множеству натуральных чисел (такие множества называют *несчетными*).

Таким образом, *счетное множество* — это бесконечное множество, элементы которого можно «перенумеровать» при помощи множества натуральных чисел, т. е. можно указать такой способ нумерации элементов множества, при котором каждый элемент множества получит свой единственный номер. Счетными множествами будут, например, множество всех четных натуральных чисел, множество всех рациональных чисел и т. д.

Приведенные примеры различных счетных множеств создают на первый взгляд несколько «парадоксальную» ситуацию: очевидно, что множество всех четных натуральных чисел является подмножеством множества всех натуральных чисел, и в то же время множество четных чисел эквивалентно множеству натуральных чисел. Этот «парадокс» кажущийся. Он связан с тем, что на практике мы обычно имеем дело с конечными множествами, т. е. с множествами, для которых изъятие хотя бы одного элемента приводит к уменьшению числа элементов множества. В результате получается новое конечное множество, которое уже не будет эквивалентно исходному, так как конечные множества эквивалентны лишь в случае равенства числа их элементов. Совсем по-другому обстоит дело со счетными множествами. Изъятие из счетного множества конечного (а иногда даже бесконечного) числа элементов по-прежнему оставляет множество счетным.

---

*) Множество всех натуральных чисел, меньших или равных некоторому натуральному числу n, называется *отрезком натурального ряда* и обозначается $|1, n|$ или $\overline{1, n}$.

В качестве примера покажем, что множество всех натуральных чисел и множество всех четных натуральных чисел эквивалентны между собой. Начертим две параллельные числовые оси Ox и $O'x'$ (рис. 1.6), выбирая единичные масштабные отрезки так, чтобы единичный отрезок оси Ox был в два раза больше единичного отрезка оси $O'x'$. Множество натуральных чисел будет изображаться точками числовых осей Ox и $O'x'$. Поставим в соответствие числу 2 числовой оси $O'x'$ число 1 числовой оси Ox, числу 4 числовой оси $O'x'$ — число 2 числовой оси Ox, числу 6 — число 3 и т. д. При этом окажется, что любому четному натуральному числу соответствует единственное натуральное число и, наоборот, любое натуральное число отвечает единственному четному натуральному числу, т. е. между множеством всех четных натуральных чисел и множеством всех

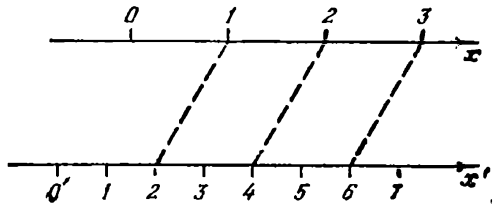


Рис. 1.6

натуральных чисел установлено взаимно однозначное соответствие, и, значит, эти множества эквивалентны.

Приведенный пример показывает, что множество натуральных чисел содержит эквивалентное ему собственное подмножество (множество четных чисел). Таким с в о й с т в о м — содержать собственное подмножество, которое эквивалентно самому множеству, обладают все бесконечные множества. Указанное свойство можно взять в качестве определения бесконечного множества.

С в о й с т в а б е с к о н е ч н ы х м н о ж е с т в:

1) всякое подмножество счетного множества конечно или счетно;

2) сумма любого конечного или счетного множества счетных множеств есть счетное множество;

3) всякое бесконечное множество содержит счетное подмножество.

Последнее свойство говорит о том, что среди бесконечных множеств счетные множества — «самые маленькие».

Существуют бесконечные множества, не являющиеся счетными; это *несчетные множества*. Заметим, что такое «определение» несчетного множества нельзя считать вполне удовлетворительным по следующей причине: вводя понятие бесконечного множества, мы были уверены, что хотя бы одно такое множество существует. Этим множеством является множество натуральных чисел. При данном же «определении» несчетного множества уже приходится в качестве теоремы доказывать, что такие множества существуют. Эта т е о р е м а формулируется так:

Множество действительных чисел, заключенных между 0 *и* 1, *несчетно.*

Мы не будем приводить доказательство этой фундаментальной теоремы теории множеств, а лишь заметим, что оно основано на представлении действительных чисел в виде бесконечных десятичных дробей. Суть же доказательства состоит в том, что, как бы мы ни нумеровали числа, лежащие между нулем и единицей, используя все натуральные числа, всегда найдется такое действительное число, большее нуля и меньшее единицы, которое будет пропущено.

Если множество эквивалентно множеству всех действительных чисел, бо́льших нуля и меньших единицы, то говорят, что оно *имеет мощность континуума.*

Примерами множеств, эквивалентных множеству действительных чисел, заключенных между нулем и единицей, т. е. имеющих мощность континуума, являются: множество всех точек любого отрезка прямой, множество всех точек прямой, множество всех прямых на плоскости и т. д.

Легко заметить, что приведенные примеры создают ту же «парадоксальную» ситуацию, которая возникала и в случае счетных множеств, а именно: множество чисел, лежащих между числами 0 и 1, оказывается эквивалентным множеству всех чисел, больших нуля и меньших, например, двух. В случае несчетных множеств «парадокс» объясняется так же, как и в случае счетных множеств, поэтому не будем на нем подробно останавливаться. Здесь мы лишь покажем, каким образо можно установить взаимно однозначное соответствие между множеством чисел, больших нуля и меньших единицы, и множеством чисел, больших нуля и меньших двух, и тем самым доказать их эквивалентность. Начертим две параллельные числовые оси Ox и $O'x'$ и выберем на них одинаковые масштабные отрезки (рис. 1.7).

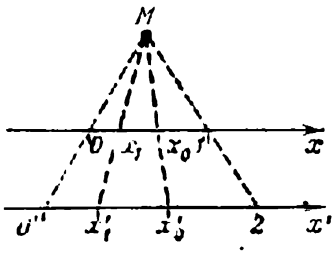


Рис. 1.7

Способ, позволяющий установить взаимно однозначное соответствие между точками, принадлежащими отрезку $[0; 1]$ оси Ox, и точками, принадлежащими отрезку $[0; 2]$ оси $O'x'$, состоит в следующем. Проведем через точки O и O' и точки 1 и 2 прямые, которые пересекутся в некоторой точке M. Возьмем любую точку $x_0 \in [0, 1]$ оси Ox и проведем прямую через точки M и x_0. Эта прямая пересечет ось $O'x'$ в некоторой точке x_0', которая и будет образом точки x_0 при отображении отрезка $[0; 1]$ оси Ox на отрезок $[0; 2]$ оси $O'x'$. Точка x_0' будет единственной точкой пересечения прямых Mx_0 и $O'x'$, так как эти прямые не параллельны. И наоборот, взяв любую точку $x_1' \in [0; 2]$ оси $O'x'$ и проведя прямую через точку M и точку x_1', принадлежашую оси $O'x'$, получим единственную точку пересечения данной прямой с осью Ox, которую обозначим x_1. Точка x_1 будет образом точки x_1' при отображении множества точек отрезка $[0; 2]$ оси $O'x'$ на множество точек отрезка $[0; 1]$ оси Ox. Таким образом, установлено взаимно однозначное соответствие между данными множествами и тем самым доказана их эквивалентность.

Приведем еще один пример множества, имеющего мощность континуума, также несколько противоречащий «здравому смыслу». На первый взгляд кажется почти очевидным, что плоский квадрат со стороной, равной единице длины, содержит внутри себя «больше» точек, чем отрезок единичной длины. Однако, как показал Г. Кантор, мощность множества точек, находящихся в квадрате, равна мощности континуума, т. е. мощности точек отрезка единичной длины. Такую же мощность имеет множество точек, находящихся внутри куба с ребром единичной длины.

Наряду с приведенными выше двумя различными мощностями (мощностью счетного множества и мощностью континуума) существуют множества, имеющие бо́льшие мощности, чем мощность континуума.

§ 3. Множества с бинарными операциями

3.1. Бинарные операции в множествах. Пусть A — некоторое (конечное или бесконечное) множество, состоящее из элементов

$$a, b, c, d, \ldots$$

Бинарной операцией (или просто *операцией*) называют отображение множества A в себя, которое каждой упорядоченной паре элементов $(a; b)$ из множества A ставит в соответствие некоторый третий элемент (*образ* пары элементов $(a; b)$) из того же множества A.

В тех случаях, когда бинарную операцию называют *умножением*, образ пары элементов $(a; b)$ называют их *произведением* и обозначают ab или $a\,b$. Если при этом $ab=ba$, то говорят, что операция умножения *коммутативна*.

В тех случаях, когда бинарную операцию называют *сложением*, образ пары элементов $(a; b)$ называют *суммой* и обозначают $a+b$. Если при этом $a+b=b+a$, то говорят, что операция сложения *коммутативна*.

Пусть a, b, c — три различных элемента множества A, в котором определена бинарная операция (скажем, умножение). Произведение ab (образ пары $(a; b)$) — элемент множества A, поэтому упорядоченной паре $(ab; c)$ будет соответствовать элемент $(ab)c$. Для той же тройки элементов a, b, c можно составить и произведение $a(bc)$ — образ пары $(a; bc)$. Если при этом $(ab)c=a(bc)$, то говорят, что бинарная операция *ассоциативна*.

Элемент $e\in A$, удовлетворяющий равенствам

$$ea=ae=a$$

для любого $a\in A$, называется *единичным элементом* множества A относительно выбранной бинарной операции. (Когда бинарная операция — сложение, то единичный элемент обычно обозначают символом 0, называют *нулевым элементом* и пишут $a+0=0+a=a$.)

3.2. Изоморфизм множеств. Пусть даны два множества A и A' и в каждом из этих множеств определено по одной (не обязательно одноименной) бинарной операции. Множества A и A' называются *изоморфными*, если существует взаимно однозначное отображение $f\colon A\to A'$, сохраняющее бинарную операцию, а именно: если элементы a' и b' из множества A' являются образами элементов a и b из множества A при отображении f, то $a'b'$ есть образ элемента ab.

Например, рассмотрим два подмножества множества натуральных чисел: множество всех четных чисел и множество всех чисел, кратных числу 5. В этих множествах определена бинарная операция (сложение): сумма любых двух четных чисел — четное число, а сумма любых двух чисел, кратных 5,—число, кратное 5. Поставим в соответствие четному числу $2n$ число $5n$ ($n\in\mathbf{N}$). Это соответствие между множеством всех четных чисел и множеством всех чисел, кратных числу 5, есть взаимно однозначное отображение, обладающее перечисленными выше свойствами, и, следовательно, данные множества изоморфны.

Изоморфные множества с бинарными операциями могут отличаться друг от друга как природой своих элементов, так и названием бинарной операции.

Например, рассмотрим два множества чисел: множество всех действительных чисел $\mathbf{R}$ и множество всех положительных действительных чисел $\mathbf{R}_+$. Поставим в соответствие положительному числу $a\in\mathbf{R}_+$ его натуральный логарифм $\ln a$, устанавливая тем самым взаимно однозначное соответствие между множеством всех действительных чисел $\mathbf{R}$ и множеством всех положительных действительных чисел $\mathbf{R}_+$.

Известное логарифмическое равенство

$$\ln(a\cdot b)=\ln a+\ln b,\qquad a,\ b\in\mathbf{R}_+,$$

показывает, что множество положительных действительных чисел с операцией умножения изоморфно множеству всех действительных чисел с операцией сложения.

Таким образом, изоморфные множества неразличимы с точки зрения свойств операций: все, что может быть доказано для одного множества с некоторой бинарной операцией на основании свойств этой операции без использования природы элементов множества, автоматически переносится на все изоморфные множества. Вследствие этого понятие изоморфизма позволяет, отвлекаясь от

природы элементов множеств, обратить основное внимание на изучение самих бинарных операций.

3.3. Группы. Непустое множество $G=\{a;\ b;\ c;\ \ldots\}$ называется *группой*, если в этом множестве определена бинарная операция (обычно называемая *умножением*) так, что:

1) бинарная операция в множестве G ассоциативна, т. е.

$$a(bc)=(ab)c;$$

2) множество G содержит такой элемент e (называемый *левой единицей*), что для каждого элемента $a\in G$ выполняется равенство

$$ea=a;$$

3) для каждого элемента $a\in G$ в множестве G существует такой элемент b, что

$$ba=e;$$

элемент b называется *левым обратным элементом* и обозначается a^{-1}:

$$a^{-1}a=e.$$

Операция в группе G не обязана быть коммутативной. Если же она коммутативна, то группа G называется *коммутативной* (или *абелевой*).

Простейшие свойства группы:

1) каждая группа G имеет единственную левую и единственную правую единицы, и эти единицы равны:

$$ea=ae=a;$$

2) каждый элемент группы G имеет единственный левый и единственный правый обратные элементы, и эти элементы равны:

$$a^{-1}a=aa^{-1}=e.$$

Из последнего свойства следуют правила, называемые *законами сокращения*: для любых трех элементов a, b, c группы G из равенств $ca=cb$, $ac=bc$ следует равенство $a=b$.

Если группа G состоит из конечного числа элементов, то она называется *конечной*, а число элементов в ней — *порядком* группы; в противном случае группа называется *бесконечной*. Бесконечная группа может быть как счетной, так и несчетной.

Если групповую бинарную операцию называют умножением (что и было принято выше), то группа называется *мультипликативной*. Если же групповую операцию называют сложением, то группа называется *аддитивной*. В этом случае вместо единицы группы говорят о *нуле* группы и обозначают его символом 0; элемент, обратный элементу a, называют *противоположным* и обозначают $-a$.

Примеры. 1. Множество всех целых чисел образует группу относительно операции сложения — аддитивную группу целых чисел. Эта группа абелева в силу коммутативности операции сложения целых чисел. Роль единичного элемента в ней играет число нуль.

Множества рациональных, действительных и комплексных чисел относительно операции сложения также образуют аддитивные абелевы группы.

2. Множество всех положительных действительных чисел относительно умножения образует абелеву группу.

3. Множество всех подстановок n-й степени (см. п. 2.2 гл. 10) с бинарной операцией (умножением) образует некоммутативную группу. Это нетрудно доказать, проверяя выполнение свойств 1) — 3), определяющих группу: произведение подстановок опять будет подстановкой; роль единицы будет играть тождественная подстановка; для любой подстановки n-й степени существует обратная подстановка; операция умножения подстановок ассоциативна. Убедимся на примере подстановок четвертой степени, что умножение подстановок

некоммутативно. Пусть a и b — две подстановки четвертой степени:

$$a=\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix}, \qquad b=\begin{pmatrix}1&2&3&4\\3&1&2&4\end{pmatrix}.$$

Тогда

$$ab=\begin{pmatrix}1&2&3&4\\1&3&4&2\end{pmatrix}, \qquad ba=\begin{pmatrix}1&2&3&4\\4&2&1&3\end{pmatrix},$$

откуда видно, что $ab \neq ba$, т. е. умножение подстановок некоммутативно.

Группа подстановок n-й степени называется *симметрической группой* степени n. Она является конечной группой порядка $n!$ и обозначается S_n.

3.4. Кольца. Множество R называется *кольцом*, если в нем определены две бинарные операции — *сложение* и *умножение*, удовлетворяющее условиям:

1) множество R — коммутативная группа по сложению, т. е.

$$a+b=b+a, \qquad a+(b+c)=(a+b)+c,$$
$$a+0=a, \qquad a+(-a)=0;$$

2) для любых $a, b, c \in R$ (ассоциативность)

$$a(bc)=(ab)c;$$

3) сложение и умножение связаны законами дистрибутивности:

$$a(b+c)=ab+ac, \qquad (b+c)a=ba+ca.$$

Множество R называется *коммутативным* кольцом, если к этим условиям добавлено условие коммутативности умножения.

Кольцом с единицей называют кольцо, содержащее такой элемент e (мультипликативную единицу), что

$$ea=a$$

для всех $a \in R$. Если R — кольцо с единицей, то может оказаться, что данный элемент a кольца R может иметь мультипликативный обратный элемент a^{-1}, а может и не иметь его.

Примеры. 1. Кольцом является множество всех целых чисел, множество всех действительных чисел, множество всех комплексных чисел.

2. Множество четных целых чисел будет кольцом без единицы, в то время как множество нечетных чисел кольцом не является, так как сумма двух нечетных чисел есть число четное (незамкнутость относительно операции сложения).

3.5. Поля. Кольцо с единицей, которое состоит не только из одного нуля и которое для каждого своего элемента a, отличного от нуля, содержит также его мультипликативный обратный элемент a^{-1}, называется *полем*.

Используя определения кольца и поля, нетрудно проверить, что:

1) множество всех натуральных чисел с определенными в нем бинарными операциями сложения и умножения не является ни кольцом, ни полем;

2) множество всех целых чисел является коммутативным кольцом с единицей;

3) множество всех рациональных чисел, полученное расширением множества целых чисел присоединением к нему частных от деления любых двух целых чисел друг на друга (за исключением деления на нуль), является полем;

4) множество всех действительных чисел, полученное расширением множества рациональных чисел присоединением к нему новых элементов — иррациональных чисел, также является полем;

5) множество комплексных чисел, полученное расширением множества действительных чисел присоединением к нему нового элемента — корня уравнения $x^2+1=0$, является полем.

Из всех числовых полей поле рациональных чисел — «самое маленькое», так как не существует числовых полей, отличных от поля рациональных чисел и целиком в нем содержащихся, и, кроме того, поле рациональных чисел содержится во всяком числовом поле.

Глава 2

ДЕЙСТВИТЕЛЬНЫЕ ЧИСЛА

Число — это важнейшее математическое понятие. Натуральные числа, используемые для счета в практической деятельности, появились на самых ранних этапах развития человеческой цивилизации. Первоначально понятие отвлеченного числа отсутствовало — число было «привязано» к пересчитываемым предметам, и в языке первобытных народов существовали различные словесные обороты для обозначения одного и того же числа разных предметов. Отвлеченное понятие натурального числа (не связанного с пересчетом конкретных предметов) появляется и закрепляется с развитием письменности и введением для обозначения чисел определенных символов.

Появление дробных (положительных рациональных) чисел было связано с необходимостью производить измерения, т. е. процедуру, в которой какая-либо величина сравнивается с другой величиной того же рода, выбираемой в качестве эталона (единицы измерения). Но так как единица измерения не всегда укладывалась целое число раз в измеряемой величине, и пренебречь этим обстоятельством в ряде случаев было нельзя, то возникла практическая потребность ввести более «мелкие» числа, нежели натуральные. Это и было источником возникновения наиболее «простых» дробей, таких, как половина, треть, четверть и т. д. Дальнейшее развитие понятия числа было обусловлено уже не только непосредственной практической деятельностью человека, но и явилось следствием развития математики.

Введение отрицательных чисел было вызвано развитием алгебры как науки, дающей общие способы решения арифметических задач независимо от их конкретного содержания и исходных числовых данных. Отрицательные числа схематически употреблялись индийскими математиками еще в VI—XI веках. В европейской науке отрицательные числа окончательно вошли в употребление лишь после работ Р. Декарта в XVII веке, давшего их геометрическое истолкование.

Множество рациональных чисел оказывается достаточным для удовлетворения большинства практических потребностей — с помощью рациональных чисел измерения можно выполнять с любой наперед заданной степенью точности.

Дальнейшее расширение понятия числа произошло в XVII веке в период зарождения современной математики, когда возникла необходимость ввести четкое определение понятия числа. Такое определение было дано одним из основоположников математического анализа И. Ньютоном в книге «Всеобщая арифметика»: «Под числом мы понимаем не столько множество единиц, сколько отвлеченное отношение какой-нибудь величины к другой величине того же рода, принятой нами за единицу». Эта формулировка дает единое определение действительного числа, как рационального, так и иррационального. (О существовании несоизмеримых отрезков, отношение которых есть

число иррациональное, было известно еще ученым Древней Греции.) В дальнейшем, в 70-х годах XIX века строгая теория действительного числа была развита в работах Р. Дедекинда, Г. Кантора и К. Вейерштрасса.

§ 1. Натуральные числа

1.1. Множество натуральных чисел. *Натуральные числа* — это числа, используемые для счета:

$$1,\ 2,\ 3,\ 4,\ \ldots,\ n,\ \ldots \qquad (1)$$

Из любых двух соседних чисел в записи (1) число, стоящее справа, называется *последующим* относительно числа, стоящего слева.

Натуральные числа (1) образуют множество, называемое *множеством натуральных чисел*. Множество всех натуральных чисел обозначается символом **N**:

$$\mathbf{N} = \{1;\ 2;\ 3;\ \ldots;\ n;\ \ldots\}.$$

Множество натуральных чисел является *упорядоченным множеством*, т. е. для любых двух натуральных чисел m и n имеет место одно из следующих соотношений:

$$\begin{aligned} &\text{либо } m = n \quad (m \text{ равно } n),\\ &\text{либо } m < n \quad (m \text{ меньше } n),\\ &\text{либо } n < m \quad (n \text{ меньше } m). \end{aligned}$$

Наименьшим натуральным числом является 1 (единица).

В множестве натуральных чисел вводятся две основные арифметические операции — сложение и умножение. Для обозначения этих операций используются соответственно символы $+$ и $\cdot$ (или $\times$).

Сложение натуральных чисел. Каждой паре натуральных чисел $(n;\ p)$ ставится в соответствие натуральное число s, называемое их *суммой*. Сумма s состоит из стольких единиц, сколько их содержится в числах n и p. О числе s говорят, что оно получено в результате сложения чисел n и p, и пишут

$$s = n + p. \qquad (2)$$

Числа n и p в записи (2) называются *слагаемыми*. Операция сложения натуральных чисел:

1) коммутативна: $n + p = p + n$;

2) ассоциативна: $(n + p) + k = n + (p + k)$.

Умножение натуральных чисел. Каждой упорядоченной паре натуральных чисел $(n;\ p)$ ставится в соответствие натуральное число m, называемое их *произведением*. Произведение m состоит из стольких единиц, сколько их содержится в числе n, взятых столько раз, сколько единиц содержится в числе p. О числе m говорят, что оно получено в результате умножения чисел n и p, и пишут

$$m = n \cdot p \quad \text{или} \quad m = n \times p. \qquad (3)$$

Числа n и p в записи (3) называются *сомножителями*.

Операция умножения натуральных чисел:

1) коммутативна: $n \cdot p = p \cdot n$;

2) ассоциативна: $(n \cdot p) \cdot k = n \cdot (p \cdot k)$.

Операции сложения и умножения натуральных чисел связаны законом дистрибутивности умножения относи-

тельно сложения:

$$(n+p)\cdot k = n\cdot k + p\cdot k.$$

Таким образом, сумма и произведение любых двух натуральных чисел опять будут натуральными числами. Поэтому говорят, что множество всех натуральных чисел замкнуто относительно операций сложения и умножения.

Вычитание натуральных чисел. Вычитание натуральных чисел есть операция, обратная сложению — упорядоченной паре натуральных чисел $(n; p)$ ставится в соответствие такое натуральное число r, что

$$n = p + r.$$

О числе r говорят, что оно получено в результате вычитания числа p из числа n, и пишут

$$r = n - p.$$

Число r называется *разностью* чисел n и p; число n называется *уменьшаемым*, а число p — *вычитаемым*.

В множестве натуральных чисел разность двух натуральных чисел $r = n - p$ существует тогда и только тогда, когда $n > p$; поэтому говорят, что множество натуральных чисел не замкнуто относительно вычитания.

Например, натуральное число 5 больше натурального числа 3. Их разность существует и равна натуральному числу 2:

$$5 - 3 = 2.$$

Натуральное число 6 меньше натурального числа 8. Их разность $6 - 8$ уже не будет натуральным числом.

Деление натуральных чисел. Деление натуральных чисел есть операция, обратная умножению — упорядоченной паре натуральных чисел $(n; p)$ ставится в соответствие такое натуральное число q, что

$$n = p\cdot q.$$

О числе q говорят, что оно получено в результате деления числа n на число p, и пишут

$$q = \frac{n}{p}, \quad \text{или} \quad q = n/p, \quad \text{или} \quad q = n:p.$$

Число q называется *частным* натуральных чисел n и p; число n называется *делимым*, а число p — *делителем*.

В множестве натуральных чисел частное определено не для любой пары натуральных чисел $(n; p)$, т. е. множество натуральных чисел не замкнуто относительно операции деления.

Например, положим $n = 7$, $p = 2$. Для этой пары натуральных чисел нельзя подобрать такое натуральное число q, чтобы выполнялось равенство

$$7 = 2\cdot q.$$

Натуральная степень числа. Свойство ассоциативности операции умножения натуральных чисел позволяет ввести понятие натуральной степени натурального числа: *n-й степенью* натурального числа m называется натуральное число k, полученное в результате умножения числа m самого на себя n раз:

$$k = \underbrace{m\cdot m\cdot m\cdot \ldots \cdot m}_{n \text{ сомножителей}}.$$

Для обозначения n-й степени числа m обычно используется запись

$$m^n,$$

в которой число m — *основание степени*, а число n — *показатель степени*.

1.2. Аксиоматическое построение множества натуральных чисел. Выше были приведены некоторые свойства натуральных чисел, а также были введены операции над натуральными числами, подчиненные правилам коммутативности, ассоциативности и дистрибутивности.

Излагая сведения о натуральных числах и действиях с ними, мы неявно обращались к интуитивному пониманию многих понятий, которыми мы ежедневно пользуемся и при этом получаем правильные результаты. Например, нам кажется естественным, что $3+2=2+3$, и мы даже не задумываемся, откуда берется это свойство операции сложения натуральных чисел. В математике, естественно, возникает вопрос — а сколько же и какие именно первичные утверждения (аксиомы) о натуральных числах необходимо выдвинуть, чтобы из этих аксиом можно было получить в виде теорем все известные из нашего жизненного опыта свойства натуральных чисел и операций над ними.

Оказывается, что все свойства натуральных чисел могут быть выведены как теоремы из пяти аксиом и формул, определяющих операции сложения и умножения натуральных чисел.

Аксиомы натуральных чисел (аксиомы Пеано).

I. 1 (единица) есть натуральное число.

II. Для каждого натурального числа n имеется только одно натуральное число, называемое его *последующим* и обозначаемое $S(n)$.

III. Всегда $S(n) \neq 1$.

IV. Из равенства $S(n)=S(m)$ следует $n=m$.

V. *Принцип полной индукции.* Множество натуральных чисел, содержащее 1 и для каждого из n элементов следующий за ним элемент $S(n)$, содержит все натуральные числа.

Сложение и умножение натуральных чисел определяются формулами

$$n+1=S(n), \qquad m+S(n)=S(m+n),$$
$$n\cdot 1=n, \qquad n\cdot S(m)=n\,m+n.$$

Из аксиом Пеано и определения операций сложения и умножения натуральных чисел как теоремы следуют законы коммутативности и ассоциативности сложения и умножения, свойство дистрибутивности умножения относительно сложения.

Из аксиом Пеано и определения операции сложения следует свойство упорядоченности множества натуральных чисел: для любых двух натуральных чисел m и n

либо $m=n$ (m равно n);

либо $m<n$ (m меньше n);

либо $n<m$ (n меньше m).

Упорядоченность множества натуральных чисел устанавливается следующим образом. Во-первых, в качестве теоремы доказывается, что для любых натуральных m и n имеет место один из следующих трех случаев:

либо $m=n$;

либо существует единственное натуральное число k, удовлетворяющее условию $n=m+k$;

либо существует единственное натуральное число p такое, что $m=n+p$.

Во-вторых, вводится определение знака $>$ (больше) и знака $<$ (меньше):

натуральное число m считают *больше* натурального числа n (и пишут $m>n$), если существует такое натуральное число k, что $m=n+k$;

натуральное число m считают *меньше* натурального числа n (и пишут $m<n$), если существует такое натуральное число p, что $m+p=n$.

Из данных определений и сформулированной выше теоремы следует свойство упорядоченности множества натуральных чисел. В силу этого свойства для натуральных чисел оказывается справедливым еще ряд утверждений, которые связаны с понятиями «больше» и «меньше». К таким утверждениям относятся, например, следующие теоремы:

1) из $m > n$ для любого k следует $m+k > n+k$;
   из $m = n$ для любого k следует $m+k = n+k$;
   из $m < n$ для любого k следует $m+k < n+k$;
2) из $m > n$ для любого k следует $m \cdot k > n \cdot k$;
   из $m = n$ для любого k следует $m \cdot k = n \cdot k$;
   из $m < n$ для любого k следует $m \cdot k < n \cdot k$;

3) В каждом непустом множестве натуральных чисел имеется наименьшее число.

1.3. Простые числа. Основная теорема арифметики. Если для двух заданных натуральных чисел n и p найдется натуральное число q такое, что $n = p \cdot q$, то говорят, что число n *делится нацело* на число p. Число p называется *делителем* числа n, а о натуральном числе n говорят, что оно *кратно* p. Если p — делитель числа n, то и q — делитель числа n.

Натуральное число, имеющее делителями только самого себя и единицу, называется *простым* числом. Остальные натуральные числа называются *составными*. Число 1 не считается простым числом.

Представление натурального числа n в виде произведения двух натуральных чисел $p \cdot q$ называется *разложением на множители*. При этом считается, что если число n простое, то оно имеет разложение на множители, состоящее из единственного числа n. Например, простое число 37 будет иметь разложение на множители, состоящее из единственного множителя 37 (а не разложение $1 \cdot 37$).

Пусть натуральное число n составное, т. е.

$$n = p \cdot q,$$

где $p \neq 1$ и $q \neq 1$. При этом возможны следующие случаи:

1) если натуральные числа p и q простые, то число n представлено в виде произведения двух простых чисел p и q;

2) если хотя бы одно из чисел p, q составное, то его (либо оба составных числа p и q) можно разложить в произведение меньших натуральных чисел, для которых возможны те же случаи. Поскольку имеется лишь конечное множество натуральных чисел, меньших n, то процесс разложения закончится через конечное число шагов. В результате получим разложение числа n на множители, каждый из которых — простое число. Представление числа n в виде произведения простых чисел называется *разложением на простые множители*.

В связи с рассмотренным разложением натурального числа на простые множители возникает вопрос — существует ли какое-нибудь другое, отличное от уже полученного, разложение натурального числа на простые множители? Ответ на этот вопрос дает следующая теорема, называемая основной теоремой арифметики.

Каждое натуральное число, отличное от 1, может быть разложено на простые множители, и притом единственным образом (если отождествлять разложения $p \cdot q$ и $q \cdot p$, где p и q — простые числа).

Объединяя в разложении числа n одинаковые простые сомножители, получаем так называемое *каноническое разложение* числа n:

$$n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdot \ldots \cdot p_s^{\alpha_s},$$

где $p_1, p_2, \ldots, p_s$ — простые, $k_1, k_2, \ldots, k_s$ — натуральные числа и все числа $p_1, p_2, \ldots, p_s$ различны.

Натуральное число называется *четным*, если среди его простых множителей есть число 2, и *нечетным*, если среди его простых множителей число 2 отсутствует.

1.4. Некоторые признаки делимости натуральных чисел.

Признак делимости на 2. Число делится на 2, если его последняя цифра есть число четное или нуль.

Признак делимости на 4. Число делится на 4, если две его последние цифры — нули или образуют число, делящееся на 4.

Признак делимости на 8. Число делится на 8, если три последние его цифры — нули или образуют число, делящееся на 8.

Признаки делимости на 3 и на 9. Число делится на 3, если сумма его цифр делится на 3. Число делится на 9, если сумма его цифр делится на 9.

Признак делимости на 5. Число делится на 5, если его последняя цифра либо нуль, либо 5.

Признак делимости на 25. Число делится на 25, если его последние две цифры — нули или образуют число, делящееся на 25.

Признак делимости на 11. Число делится на 11, если сумма цифр, стоящих на четных местах, либо равна сумме цифр, стоящих на нечетных местах, либо отличается от нее на число, делящееся на 11.

1.5. Наименьшее общее кратное. Наибольший общий делитель. Алгорим Евклида. *Общим кратным* нескольких натуральных чисел называется натуральное число, являющееся кратным для каждого из них. Наименьшее из них называется *наименьшим общим кратным* (сокращенно НОК) *).

Для того чтобы найти наименьшее общее кратное нескольких чисел, необходимо:

1) выписать канонические разложения данных чисел;

2) перечислить все простые множители, входящие хотя бы в одно из канонических разложений данных чисел;

3) возвести каждый из перечисленных простых множителей в наибольшую степень, с которой этот простой множитель входит в канонические разложения данных чисел.

Произведение полученных степеней простых множителей и даст число, являющееся наименьшим общим кратным данных чисел.

Пример. Найти НОК чисел 49 896 и 26 460.

1) Пишем канонические разложения данных чисел:

$$49\,896 = 2^3 \cdot 3^4 \cdot 7 \cdot 11, \qquad 26\,460 = 2^2 \cdot 3^3 \cdot 5 \cdot 7^2.$$

2) Выписываем простые множители, входящие хотя бы в одно каноническое разложение:

$$2;\ 3;\ 5;\ 7;\ 11.$$

3) [Наибольшая степень, с которой множитель 2 входит в канонические разложения данных чисел, равна 3; пишем число 2^3. Наибольшая степень, с которой множитель 3 входит в канонические разложения, равна 4; пишем число 3^4. Наибольшие степени, с которыми множители 5, 7 и 11 входят в канонические разложения данных чисел, равны соответственно 1, 2 и 1; пишем числа $5^1 = 5$, 7^2 и $11^1 = 11$.

*) Наименьшее общее кратное двух чисел m и n также обозначается с помощью фигурных скобок: $\{m, n\}$.

Произведение $2^3 \cdot 3^4 \cdot 5 \cdot 7^2 \cdot 11 = 1\,746\,360$ является наименьшим общим кратным данных чисел.

Общим делителем нескольких натуральных чисел называется число, являющееся делителем каждого из данных чисел. В случае, если общих делителей несколько, наибольший из них называется *наибольшим общим делителем* (сокращено НОД) *).

Если наибольший общий делитель нескольких натуральных чисел равен единице, то эти числа называются *взаимно простыми*.

Для нахождения наибольшего общего делителя нескольких чисел необходимо:

1) выписать канонические разложения данных чисел;

2) перечислить все общие простые множители, входящие в канонические разложения каждого из данных чисел;

3) возвести каждый из перечисленных простых множителей в наименьшую степень, с которой этот простой множитель входит в канонические разложения данных чисел.

Произведение полученных степеней простых множителей и будет наибольшим общим делителем данных чисел.

Пример. Найти НОД чисел 49 896 и 26 460.

1) Пишем канонические разложения данных чисел:

$$49\,896 = 2^3 \cdot 3^4 \cdot 7 \cdot 11, \qquad 26\,460 = 2^2 \cdot 3^3 \cdot 5 \cdot 7^2.$$

2) Выписываем простые множители, входящие в оба канонических разложения:

$$2;\ 3;\ 7.$$

3) Наименьшая степень, с которой множитель 2 входит в канонические разложения данных чисел, равна 2; пишем число 2^2. Наименьшая степень, с которой множитель 3 входит в канонические разложения, равна 3; пишем число 3^3. Наименьшая степень, с которой число 7 входит в канонические разложения, равна 1; пишем число $7^1 = 7$.

Произведение $2^2 \cdot 3^3 \cdot 7 = 756$ является наибольшим общим делителем данных чисел.

Свойство НОД. Наибольший общий делитель двух чисел делится на любой общий делитель этих чисел.

Наибольший общий делитель (m, n) двух чисел m и n и наименьшее общее кратное $\{m, n\}$ связаны равенством

$$m \cdot n = (m, n) \cdot \{m, n\}. \tag{4}$$

Алгоритм Евклида нахождения НОД. Наибольший общий делитель и наименьшее общее кратное двух натуральных чисел могут быть найдены, если известны канонические разложения этих чисел. Однако выписать канонические разложения достаточно больших чисел часто бывает довольно трудно. Кроме способа нахождения НОД двух чисел, основанного на канонических разложениях этих чисел, существует способ нахождения НОД, не требующий знания всех простых множителей этих чисел. Этот способ называется *алгоритмом Евклида*. Заметим, что, отыскав НОД двух чисел с помощью алгоритма Евклида, с помощью формулы (4) можно найти НОК этих чисел.

Прежде чем перейти к изложению алгоритма Евклида, введем операцию деления с остатком.

Разделить натуральное число m на натуральное число p $(m \geqslant p)$ с остатком – это значит найти такое натуральное число k и такое неотрицательное целое

*) Наибольший общий делитель двух чисел m, n также обозначается с помощью круглых скобок: (m, n).

число $r < p$, что

$$m = p \cdot k + r.$$

Числа k и r называются соответственно *частным* и *остатком* от деления числа m на число p. В случае, если $r = 0$, говорят, что остаток от деления равен нулю или что число m *делится* на число p.

Алгоритм Евклида. Пусть m и p — два натуральных числа и пусть $m > p$. Обозначим через m_1 и r_1 соответственно частное и остаток от деления m на p:

$$m = p\, m_1 + r_1, \qquad 0 \leqslant r_1 < p. \tag{5}$$

Если $r_1 > 0$, то делим p на r_1, обозначив через m_2 и r_2 соответственно частное и остаток от деления:

$$p = r_1 \cdot m_2 + r_2, \qquad 0 < r_2 < r_1. \tag{6}$$

Если r_2 все еще не нуль, делим r_1 на r_2 и получаем аналогичным образом

$$r_1 = r_2 \cdot m_3 + r_3, \qquad 0 \leqslant r_3 < r_2. \tag{7}$$

Поскольку p, r_1, r_2, ... — натуральные числа и $p > r_1 > r_2 > r_3 > \ldots$, то процесс деления после конечного числа шагов должен оборваться, т. е. мы должны притти к остатку, равному нулю. Пусть таким остатком будет $r_{n+1} = 0$, так что

$$r_{n-1} = r_n \cdot m_{n+1}. \tag{8}$$

Натуральное число r_n и будет наибольшим общим делителем чисел m и p. Чтобы убедиться в этом, покажем, что m и p делятся на r_n. В силу равенства (8) r_{n-1} делится на r_n. Тогда из равенства

$$r_{n-2} = r_{n-1} \cdot m_n + r_n,$$

предшествующего равенству (8), следует, что и r_{n-2} делится на r_n. Продвигаясь далее по цепочке построенных равенств (снизу вверх), убеждаемся, что на r_n делятся и r_{n-3}, r_{n-4}, ..., r_2, r_1, а в силу равенств (6) и (5) также и числа m и p.

Остается убедиться в том, что всякий общий делитель k чисел m и p будет делителем и числа r_n (этим, очевидно, и будет установлено, что r_n — наибольший общий делитель чисел m и p. Для этого снова пройдем всю цепочку равенств, но на этот раз сверху вниз. Предварительно перепишем цепочку равенств алгоритма Евклида в другом виде:

$$\begin{aligned} r_1 &= m - p m_1, \\ r_2 &= p - r_1 m_2, \\ r_3 &= r_1 - r_2 m_3, \\ &\cdots\cdots\cdots\cdots \\ r_n &= r_{n-2} - r_{n-1} m_n. \end{aligned}$$

Из первого равенства следует, что r_1 делится на общий делитель k чисел m и p. Но из второго равенства следует, что r_2 делится на k (так как r_1 делится на k и p делится на k). Продолжая этот процесс, приходим к тому, что r_{n-2} и r_{n-1} делятся на k. Но тогда в силу последнего равенства и r_n делится на k.

Способ отыскания наибольшего общего делителя двух чисел с помощью алгоритма Евклида в большинстве случаев оказывается самым коротким и потому практически наиболее выгодным.

Заметим, что в приведенном рассуждении попутно доказано сформулированное выше свойство НОД.

Пример. С помощью алгоритма Евклида найти НОД чисел 13 172 и 261.

Разделим число 13 172 на 261:

$$\begin{array}{r|l} 13172 & 261 \\ \hline 1305 & 50 \\ \hline 122 & \end{array}$$

Делим в согласии с алгоритмом Евклида число 261 на остаток от деления (число 122):

$$\begin{array}{r|l} 261 & 122 \\ \hline 244 & 2 \\ \hline 17 & \end{array}$$

Делим число 122 на остаток (число 17):

$$\begin{array}{r|l} 122 & 17 \\ \hline 119 & 7 \\ \hline 3 & \end{array}$$

Продолжая и далее последовательно делить каждый предыдущий остаток на каждый последующий остаток, в результате получим остаток, равный нулю:

$$\begin{array}{r|l} 17 & 3 \\ \hline 15 & 5 \\ \hline 2 & \end{array} \qquad \begin{array}{r|l} 3 & 2 \\ \hline 2 & 1 \\ \hline 1 & \end{array} \qquad \begin{array}{r|l} 1 & 1 \\ \hline 1 & 1 \\ \hline 0 & \end{array}$$

Предпоследний остаток был равен единице. Следовательно, единица и есть НОД двух данных чисел, т. е. числа 13 172 и 261 взаимно простые.

§ 2. Целые числа

2.1. Множество целых чисел. Множество целых чисел есть множество, полученное в результате добавления к множеству натуральных чисел новых объектов (которое далее также будем называть числами) — числа нуль и отрицательных целых чисел.

Число *нуль*, обозначаемое символом 0, и *отрицательные целые числа* вводятся следующим образом.

Сумма любого натурального числа n и числа 0 есть число n:

$$n+0=n.$$

Любому натуральному числу n соответствует единственное отрицательное целое число $-n$: такое, что сумма чисел n и $-n$ равна нулю:

$$n+(-n)=0.$$

Число $-n$ называется *противоположным* числу n. Число, противоположное числу $-n$, есть число n: $-(-n)=n$. Натуральные числа в множестве целых чисел называются *положительными целыми числами* *).

Множество целых чисел является упорядоченным множеством, т. е. для любых двух целых чисел m и n справедливо одно и только одно из следующих соотношений:

$$\text{либо } m=n, \quad \text{либо } m<n, \quad \text{либо } n<m.$$

Для положительных чисел n пишут $n>0$, для отрицательных чисел пишут $n<0$. Если хотят указать, что число может быть положительным или нулем, пишут $n\geqslant 0$ и говорят, что оно неотрицательно; аналогично запись $n\leqslant 0$ означает, что n отрицательно или равно нулю.

Множество всех целых чисел обозначается $\mathbf{Z}$.

2.2. Арифметические операции с целыми числами. *Абсолютным значением* (или *модулем*) числа n называется число, обозначаемое $|n|$ и вычисляемое по правилу

$$|n|=\begin{cases} n, & \text{если } n>0, \\ 0, & \text{если } n=0, \\ -n, & \text{если } n<0. \end{cases}$$

*) Далее целые числа будем обозначать малыми латинскими буквами m, n, k, ...

Абсолютное значение числа n положительно как для положительных, так и для отрицательных n и равно нулю только при $n=0$.

Пример. Найти абсолютные значения чисел 4 и -3. Так как число 4 натуральное, то $|4|=4$. Так как -3 — отрицательное число, то по правилу вычисления абсолютного значения числа $|-3|=-(-3)=3$.

Сложение целых чисел. *Суммой* двух целых чисел n и p называется целое число s, вычисляемое по правилу:

если $n>0$ и $p>0$, то $s=n+p$;

если $n<0$ и $p<0$, то $s=-(|n|+|p|)$;

если $n>0$, а $p<0$ и $|n|>|p|$, то $s=|n|-|p|$;

если $n>0$, а $p<0$ и $|n|=|p|$, то $s=0$;

если $n>0$, а $p<0$ и $|n|<|p|$, то $s=-(|p|-|n|)$;

если $n<0$, $p>0$ и $|n|>|p|$, то $s=-(|n|-|p|)$;

если $n<0$, $p>0$ и $|n|=|p|$, то $s=0$;

если $n<0$, $p>0$ и $|n|<|p|$, то $s=|p|-|n|$;

если $n=0$, то $s=p$;

если $p=0$, то $s=n$.

Сумма s двух целых чисел n и p записывается с помощью символа $+$:

$$s=n+p.$$

Пример. Вычислить сумму целых чисел 6 и -8. Данная пара чисел удовлетворяет пятому условию правила сложения двух целых чисел ($6>0$, $-8<0$, $|6|<|-8|$). Сумма этих чисел равна

$$s=-(|-8|-|6|)=-(8-6)=-2.$$

Сложение целых чисел, как и сложение натуральных чисел, коммутативно и ассоциативно.

Умножение целых чисел. *Произведением* двух целых чисел n и p называется целое число m, вычисляемое по правилу:

если $n>0$ и $p>0$, то $m=n\cdot p$;

если $n<0$ и $p<0$, то $m=|n|\cdot|p|$;

если $n<0$, а $p>0$ или если $n>0$, а $p<0$, то $m=-(|n|\cdot|p|)$;

если $n=0$ или $p=0$, то $m=0$.

Произведение двух целых чисел n и p записывается с помощью символа $\cdot$ или $\times$:

$$m=n\cdot p \qquad (\text{или } n\times p).$$

Пример. Вычислить произведение целых чисел -2 и -7. Данная пара удовлетворяет второму условию правила умножения двух целых чисел. Абсолютные значения этих чисел равны 2 и 7, а их произведение равно 14.

Умножение целых чисел, как и умножение натуральных чисел, коммутативно и ассоциативно.

Кроме того, операции сложения и умножения целых чисел, как и в случае натуральных чисел, связаны законом дистрибутивности умножения относительно сложения.

Вычитание целых чисел. *Разностью* двух целых чисел n и p называется целое число r, вычисляемое по правилу

$$r=n+(-p), \tag{1}$$

т. е. разность двух целых чисел n и p есть сумма целого числа n и числа $(-p)$, противоположного числу p. Следовательно, разность вычисляется по правилу вычисления суммы двух целых чисел. О числе r говорят, что оно получено в результате вычитания числа p из числа n, и пишут

$$r = n - p. \tag{2}$$

В записи (2) число n называется *уменьшаемым*, а число p — *вычитаемым*.

Пример. Вычислить разность двух целых чисел 2 и -7. Число, противоположное числу -7, равно 7, а потому по правилу (1) разность этих чисел равна $r = 2 + 7 = 9$.

Таким образом, множество целых чисел *замкнуто* относительно операций сложения, умножения и вычитания, т. е. для любых двух целых чисел существует единственное целое число, являющееся суммой двух заданных целых чисел; существует единственное целое число, являющееся их разностью, и, наконец, единственное целое число, являющееся их произведением.

Деление целых чисел. *Частным* от деления целого числа m на целое число n называется целое число p, которое удовлетворяет равенству

$$m = n \cdot p. \tag{3}$$

О числе p говорят, что оно получено в результате деления числа m на число n, и пишут

$$p = m : n, \qquad \text{или} \qquad p = \frac{m}{n}, \quad \text{или} \quad p = m/n.$$

В множестве целых чисел, как и в множестве натуральных чисел, операция деления не всегда выполнима — не для любой пары целых чисел m и n существует их частное. Поэтому говорят, что множество целых чисел не замкнуто относительно операции деления. Однако между операциями деления в множестве натуральных чисел и в множестве целых чисел есть одно существенное различие: если в множестве натуральных чисел частное двух натуральных чисел существовало, то оно было единственным. По-другому обстоит дело в множестве целых чисел: пусть m — произвольное целое число, а $n = 0$; тогда равенство (3) примет вид

$$m = 0 \cdot p. \tag{4}$$

Попытаемся найти такое число p, которое удовлетворяло бы равенству (4). Здесь могут быть две возможности:

если $m \neq 0$, то не существует такого целого числа p, при котором это равенство выполняется;

если $m = 0$, то p может быть любым целым числом.

Таким образом, частное от деления целого числа на число нуль либо не существует, либо определяется не единственным образом. Чтобы избежать такой неопределенности, необходимо запретить деление целого числа на число нуль.

2.3. Множество целых чисел как множество упорядоченных пар натуральных чисел. Множество натуральных чисел N замкнуто относительно сложения и умножения, но не замкнуто относительно вычитания: сумма и произведение любых двух натуральных чисел есть число натуральное, в то время как разность $n - p$ двух натуральных чисел в множестве натуральных чисел определена

только тогда, когда $n > p$. Множество целых чисел $\mathbf{Z}$ получается расширением множества натуральных чисел путем добавления новых числовых объектов, таких, что в расширенном множестве $\mathbf{Z}$:

1) множество $\mathbf{N}$ является собственным подмножеством;

2) сложение и умножение натуральных чисел в множестве $\mathbf{Z}$ совпадает с одноименными операциями в множестве $\mathbf{N}$;

3) вычитание в множестве $\mathbf{Z}$ всегда выполнимо, т. е. разность любых двух элементов (чисел) из $\mathbf{Z}$ является элементом (числом) из $\mathbf{Z}$;

4) расширенное множество $\mathbf{Z}$ минимально в том смысле, что оно не содержит собственного подмножества, удовлетворяющего условиям 1) —3).

Один из методов пополнения множества натуральных чисел до множества целых чисел состоит в следующем.

Рассмотрим множество всевозможных упорядоченных пар натуральных чисел $(m;\ n)$, для которых определены приведенные ниже правила сравнения и арифметические операции.

Законы сравнения. Пара $(m;\ n)$ считается *больше* $(>)$, *равной* $(=)$ или *меньше* $(<)$ пары $(p;\ q)$ тогда и только тогда, когда

$$m+q > n+p, \quad m+q = n+p, \quad m+q < n+p$$

соответственно.

Из определения равенства пар следует, что всякая пара переходит в равную ей, если к обоим ее элементам прибавить одно и то же натуральное число:

$$(m+r;\ \ n+r)=(m;\ n).$$

Арифметические операции

Сложение пар. *Суммой* двух пар $(m;\ n)$ и $(p;\ q)$ называется пара $(m+p;\ n+q)$. Говорят, что пара $(m+p;\ n+q)$ получена в результате сложения пар $(m;\ n)$ и $(p;\ q)$ и пишут

$$(m;\ n)+(p;\ q)=(m+p;\ n+q).$$

Ассоциативность и коммутативность операции сложения пар следует из определения суммы пар и ассоциативности и коммутативности операции сложения натуральных чисел.

Умножение пар. *Произведением* двух пар $(m;\ n)$ и $(p;\ q)$ называется пара

$$(m\cdot p+n\cdot q;\ \ m\cdot q+n\cdot p).$$

Об этой паре говорят, что она получена в результате умножения $(m;\ n)$ и $(p;\ q)$, и пишут

$$(m;\ n)\cdot(p;\ q)=(m\cdot p+n\cdot q;\ \ m\cdot q+n\cdot p).$$

Ассоциативность и коммутативность операции умножения пар, а также свойство дистрибутивности операции умножения относительно сложения следуют из определения суммы и произведения пар и соответствующих свойств операций сложения и умножения натуральных чисел.

Вычитание пар. *Разностью* пар $(m;\ n)$ и $(p;\ q)$ называется пара $(x;\ y)$, удовлетворяющая равенству

$$(m;\ n)=(x;\ y)+(p;\ q). \tag{1}$$

О паре $(x;\ y)$ говорят, что она получена в результате вычитания пары $(p;\ q)$ из пары $(m;\ n)$, и пишут

$$(m;\ n)-(p;\ q)=(x;\ y).$$

Из равенства (1), согласно определению операции сложения пар и сформулированному выше условию равенства пар, последовательно получаем

$$(m;\ n)=(x+p;\ y+q), \quad m+y+q=n+x+p.$$

Одним из решений этого уравнения с двумя неизвестными будет пара натуральных чисел

$$x=m+q;\ \ y=n+p.$$

Пара $(m+q;\ n+p)$ есть разность пар $(m;\ n)$ и $(p;\ q)$:

$$(m;\ n)-(p;\ q)=(m+q;\ \ n+p),$$

или, точнее говоря, пара $(m+q;\ n+p)$ есть одна из множества равных между собой пар, являющихся разностью пар $(m;\ n)$ и $(p;\ q)$.

Пара $(m+q;\ n+p)$ может быть записана как сумма пар $(m;\ n)+(q;\ p)$, т. е. разность пар $(m;\ n)-(p;\ q)$ может быть представлена как сумма $(m;\ n)+(q;\ p)$.

Пару $(q;\ p)$ называют *противоположной* паре $(p;\ q)$ и пишут

$$(q;\ p)=-(p;\ q).$$

Из определения разности пар также следует, что

$$(m;\ n)-(m;\ n)=(m+n;\ m+n).$$

Пару с равными элементами называют *нулем* и обозначают символом 0.

Множество натуральных чисел включается в множество упорядоченных пар следующим определением: любая пара $(m;\ n)$ при $m>n$ отождествляется с натуральным числом $m-n$.

Нетрудно проверить, что пара $(m;\ n)$, отождествляемая при $m>n$ с натуральным числом $m-n$, больше нуля (нулевой пары). Нетрудно также проверить, что при таком способе отождествления пары $(m;\ n)$ с натуральным числом $m-n$ результаты операций сложения и умножения пар $(m;\ n)$ и $(p;\ q)$ при $m>n$, $p>q$ совпадут с результатами операций сложения и умножения натуральных чисел $r=m-n$ и $s=p-q$.

Действительно, по определению суммы и произведения пар имеем

$$(m;\ n)+(p;\ q)=(m+p;\ n+q),$$
$$(m;\ n)\cdot(p;\ q)\ =(m\cdot p+n\cdot q;\ m\cdot q+n\ p).$$

Так как по условию $m>n$ и $p>q$, то

$$m+p>n+q \quad \text{и} \quad mp+nq>mq+np,$$

т. е. первые элементы пар, стоящих в правых частях равенств, больше вторых элементов, и, следовательно, эти пары можно отождествить с натуральными числами $m+p-n-q$ и $mp+nq-mq-np$ соответственно.

С другой стороны, складывая и перемножая натуральные числа $r=m-n$ и $s=p-q$, мы снова получаем те же натуральные числа $m+p-n-q$ и $mp+nq-mq-np$.

Таким образом, мы проверили, что результаты сложения и умножения пар совпадают с результатами тех же операций с натуральными числами, соответствующими данным парам.

Из определения операции вычитания пар следует, что в множестве упорядоченных пар разность двух натуральных чисел $r=m-n$ $(m>n)$ и $s=p-q$ $(p>q)$, рассматриваемых как упорядоченные пары $(m;\ n)$ и $(p;\ q)$, всегда существует, т. е. всегда найдется пара $(x;\ y)$, которая является разностью пар $(m;\ n)$ и $(p;\ q)$. При этом очевидно, что у пары $(x;\ y)$ первый элемент может быть как больше, так и меньше или равен второму, т. е. пара $(x;\ y)$ не всегда отождествима с некоторым натуральным числом.

Итак, определив указанным выше способом операции сложения, умножения и вычитания пар и отождествив пару натуральных чисел $(m;\ n)$ при $m>n$ с натуральным числом $m-n$, мы получили, что сумма и произведение двух пар, каждую из которых можно отождествить с некоторым натуральным числом, всегда существуют и могут быть отождествлены с некоторыми натуральными числами, а разность двух пар, каждую из которых можно отождествить с натуральным числом, всегда существует, но не всегда может быть отождествлена с каким-либо натуральным числом.

Обозначая пару $(x;\ y)$ при $x\leqslant y$ так же, как и при $x>y$, т. е. через $x-y$, нетрудно проверить, что пара $(x;\ y)$ при $x<y$ меньше нулевой пары

$$x-y<0$$

и противоположна паре $(y;\ x)$, отождествляемой с натуральным числом $y-x$.

Множество упорядоченных пар натуральных чисел $(m; n)$, обозначаемых $m-n$ и подчиненных перечисленным выше законам сравнения и законам арифметических операций, называют *множеством целых чисел*.

При этом числа $m-n$ при $m>n$ называют *положительными целыми числами*, число $m-n$ при $m=n$ называют *нулем*, а числа $m-n$ при $m<n$ — *отрицательными целыми числами*.

§ 3. Рациональные числа

3.1. Рациональные дроби. Рациональные дроби появились как форма записи чисел, более «мелких», нежели натуральные. История введения рациональных дробей отражена и в самих обиходных названиях некоторых из них. Например, дробь 1/2 (одну вторую) называют «половина», дробь 1/4 (одну четвертую) называют «четверть» и т. д. Иногда дают интуитивно понятное, но математически нестрогое определение рациональной дроби как некоторой части единицы или нескольких равных частей единицы. При таком «определении» понятия рационального числа и рациональной дроби оказываются отождествленными. Математически более строго множество рациональных дробей может быть определено следующим образом.

Рациональными дробями называют упорядоченные пары целых чисел $(m; n)$ (из которых второе отлично от нуля), подчиненные перечисленным ниже правилам сравнения и законам арифметических действий.

Первое число пары $(m; n)$ называют *числителем* дроби, второе — *знаменателем*. Рациональную дробь $(m; n)$ обычно записывают в виде m/n.

Законы сравнения. Две пары $(m; n)$ и $(p; q)$ называют *эквивалентными* и пишут

$$(m; n) \sim (p; q) \qquad \left(\text{или } \frac{m}{n} \sim \frac{p}{q}\right),$$

если $m \cdot q = n \cdot p$.

Из этого определения следует, что всякая пара $(m; n)$ переходит в эквивалентную, если оба ее элемента умножить на одно и то же отличное от нуля число, т. е.

$$(m \cdot k; n \cdot k) \sim (m; n) \qquad \left(\text{или } \frac{m \cdot k}{n \cdot k} \sim \frac{m}{n}\right).$$

Переход от дроби $\frac{m \cdot k}{n \cdot k}$ к эквивалентной дроби $\frac{m}{n}$ называют *сокращением* дроби на число k. Используя правило сокращения дробей, любую рациональную дробь можно привести к эквивалентной рациональной дроби, знаменатель которой — натуральное число. Остальные законы сравнения рациональных дробей удобно формулировать, считая, что знаменатели дробей — натуральные числа.

Пара $(m; n)$ считается *больше* пары $(p; q)$:

$$(m; n) > (p; q) \qquad \left(\frac{m}{n} > \frac{p}{q}\right),$$

если $m \cdot q > n \cdot p$.

Пара $(m; n)$ считается *меньше* пары $(p; q)$:

$$(m; n) < (p; q) \qquad \left(\frac{m}{n} < \frac{p}{q}\right),$$

если $m \cdot q < n \cdot p$.

Арифметические операции.

Сложение. *Суммой* двух пар $(m; n)$ и $(p; q)$ называется пара

$$(m \cdot q + n \cdot p;\ n \cdot q).$$

Говорят, что пара $(m \cdot q + n \cdot p;\ n \cdot q)$ получена в результате *сложения* пар $(m; n)$ и $(p; q)$, и пишут

$$(m; n) + (p; q) \sim (m \cdot q + n \cdot p;\ n \cdot q)$$
$$\left(\text{или } \frac{m}{n} + \frac{p}{q} \sim \frac{m \cdot q + n \cdot p}{nq}\right).$$

Ассоциативность и коммутативность операции сложения пар следует из определения суммы пар и из ассоциативности и коммутативности операций сложения целых чисел.

В результате сложения любой пары $(m; n)$ с парой $(0; q)$ будет получаться пара $(m; n)$. Пару $(0; q)$ называют *нулевой* парой.

Вычитание. *Разностью* пар $(m; n)$ и $(p; q)$ называется пара $(x; y)$ такая, что

$$(x; y) + (p; q) \sim (m; n).$$

Используя определение суммы пар и определение эквивалентных пар, последовательно получаем

$$(x \cdot q + y \cdot p;\ yq) \sim (m; n),$$
$$x \cdot q \cdot n + y \cdot p \cdot n = y \cdot q \cdot m.$$

Одно из решений этого уравнения с двумя неизвестными есть пара целых чисел

$$x = m \cdot q - n \cdot p, \quad y = n \cdot q.$$

О паре $(m \cdot q - n \cdot p;\ n \cdot q)$ говорят, что она получена в результате *вычитания* пары $(p; q)$ из пары $(m; n)$, и пишут

$$(m; n) - (p; q) \sim (m \cdot q - n \cdot p;\ n \cdot q) \quad \left(\text{или } \frac{m}{n} - \frac{p}{q} \sim \frac{m \cdot q - n \cdot p}{n \cdot q}\right).$$

Если из некоторой пары $(m; n)$ вычесть эквивалентную ей пару $(k \cdot m;\ k \cdot n)$, то их разность будет нулевой парой $(0; n)$.

Из определения разности пар $(m; n)$ и $(p; q)$ следует, что

$$(m; n) - (p; q) \sim (m \cdot q - n \cdot p;\ n \cdot q) \sim (m; n) + (-p; q),$$

т. е. вычитание пары $(p; q)$ из пары $(m; n)$ может быть заменено сложением пары $(m; n)$ с парой $(-p; q)$, которая называется *противоположной* паре $(p; q)$.

Умножение. *Произведением* двух пар $(m; n)$ и $(p; q)$ называется пара $(m \cdot p;\ n \cdot q)$. Говорят, что пара $(m \cdot p;\ n \cdot q)$ получена в результате *умножения* пар $(m; n)$ и $(p; q)$, и пишут

$$(m; n) \cdot (p; q) \sim (m \cdot p;\ n \cdot q) \qquad \left(\text{или } \frac{m}{n} \cdot \frac{p}{q} \sim \frac{m \cdot p}{n \cdot q}\right).$$

Ассоциативность и коммутативность операции умножения пар, а также дистрибутивность умножения относительно сложения следуют

из определения произведения и суммы пар и соответствующих свойств операций умножения и сложения целых чисел.

Деление. *Частным* двух пар $(m; n)$ и $(p; q)$ при $p \neq 0$ называется такая пара $(x; y)$, что

$$(p; q)\cdot(x; y) \sim (m; n).$$

Используя определение произведения пар и определение эквивалентных пар, последовательно получаем

$$(px; qy) \sim (m; n),$$
$$p\cdot n\cdot x = q\cdot m\cdot y.$$

Одно из решений этого уравнения с двумя неизвестными есть пара целых чисел

$$x = q\cdot m, \quad y = p\cdot n.$$

О паре $(q\cdot m; p\cdot n)$ говорят, что она получена в результате *деления* пары $(m; n)$ на пару $(p; q)$, и пишут

$$(m; n):(p; q) \sim (q\cdot m; p\cdot n) \qquad \left(\text{или } \frac{m}{n}:\frac{p}{q} \sim \frac{m\cdot q}{n\cdot p}\right).$$

Из определения частного пар $(m; n)$ и $(p; q)$ следует, что

$$(m; n):(p; q) \sim (q\cdot m; p\cdot n) \sim (m; n)\cdot(q; p),$$

т. е. деление пары $(m; n)$ на пару $(p; q)$ может быть заменено умножением пары $(m; n)$ на пару $(q; p)$, которая называется *обратной* паре $(p; q)$.

Согласно правилам сравнения дробей пара $(m; n)$ (n — натуральное число) больше нуля (нулевой пары $(0; q)$), если $m > 0$, и меньше нуля (нулевой пары $(0; q)$), если $m < 0$.

Рациональную дробь m/n, где m — целое число, а n — натуральное, называют *положительной* и пишут $m/n > 0$, если m положительно, и *отрицательной*: $m/n < 0$, если m отрицательно.

Отрицательную рациональную дробь m/n (m — отрицательное целое число, n — натуральное) обычно записывают в виде $-\frac{|m|}{n}$. Например, отрицательную рациональную дробь $\frac{-5}{7}$ записывают в виде $-\frac{5}{7}$.

Положительная рациональная дробь m/n называется *правильной*, если ее числитель меньше знаменателя ($m < n$), и *неправильной*, если ее числитель больше или равен знаменателю ($m \geqslant n$). Если положительная рациональная дробь неправильна, ее числитель представим в виде $m = n\cdot k + r$, где k — натуральное число, а r — неотрицательное целое число, удовлетворяющее условию $r < n$. Число k называется *целой частью* дроби.

Если $r \neq 0$, то неправильную дробь m/n иногда записывают в виде

$$k\frac{r}{n}. \qquad (1)$$

Запись (1) называют записью неправильной дроби в виде *смешанной* дроби. Например, неправильная рациональная дробь 10/3 может быть записана в виде смешанной дроби $3\frac{1}{3}$.

Дробь называется *несократимой*, если ее числитель и знаменатель — взаимно простые числа. Всякую рациональную дробь можно привести к эквивалентной ей несократимой дроби.

Пример. Привести рациональную дробь 15/75 к эквивалентной ей несократимой дроби.

Разложим числа 15 и 75 в произведение простых сомножителей:

$$15=3\cdot 5.\quad 75=3\cdot 5\cdot 5.$$

Дробь $\frac{15}{75}$ может быть записана в виде $\frac{3\cdot 5}{3\cdot 5\cdot 5}$. Сокращая одинаковые сомножители в числителе и знаменателе, приводим дробь 15/75 к эквивалентной несократимой дроби 1/5.

3.2. **Рациональные числа.** *Рациональным числом* называется множество всех эквивалентных между собой рациональных дробей. В согласии с этим определением различные эквивалентные между собой дроби — это лишь различные записи одного и того же рационального числа. Так, три различные эквивалентные рациональные дроби $\frac{2}{3}$; $\frac{4}{6}$; $\frac{-6}{-9}$ — это различные записи одного и того же рационального числа.

Можно дать и другое определение рационального числа, отождествляя его не с множеством всех рациональных эквивалентных между собой дробей, а с некоторой фиксированной дробью этого множества. Один из возможных способов выделения такой фиксированной дроби заключается в следующем.

Возьмем любую дробь m/n из множества эквивалентных между собой дробей (здесь m — целое число, отличное от нуля, n — натуральное). Если числа $|m|$ и n взаимно простые, то считаем, что дробь m/n и есть нужная фиксированная дробь. Если же числа $|m|$ и n не являются взаимно простыми, то делим числитель и знаменатель дроби на наибольший общий делитель чисел $|m|$ и n. Получаем дробь m_1/n_1, в которой $|m_1|$ и n_1 — взаимно простые числа. Найденная дробь m_1/n_1 и есть нужная фиксированная дробь.

Теперь можно дать следующее определение рационального (отличного от нуля) числа.

Рациональное число — это число, которое может быть записано в виде m/n, где $|m|$ и n — взаимно простые (несократимые) натуральные числа.

Запись рационального числа в виде m/n, где $|m|$ и n — взаимно простые числа, называется записью рационального числа в виде *несократимой* дроби. Рациональное число m/n будет *целым*, если знаменатель несократимой рациональной дроби равен единице.

Сравнение рациональных чисел, понятия суммы, произведения, разности и частного двух рациональных чисел вводятся так же, как и соответствующие понятия для рациональных дробей. Математическая запись сравнения рациональных чисел, операций сложения, умножения, вычитания и деления двух рациональных чисел — такая же, как и запись соответствующих операций над рациональными дробями, лишь вместо знака эквивалентности следует ставить знак равенства.

Множество рациональных чисел обозначается Q.

Для вычисления суммы, разности, произведения и частного рационального числа и целого числа n достаточно записать целое число n в виде дроби, знаменатель которой равен единице, т. е. в виде $n/1$,

и воспользоваться правилами вычисления суммы, разности, произведения и частного двух рациональных дробей.

Примеры. 1. Умножить целое число 3 на рациональное число 5/7. Запишем целое число 3 в виде дроби со знаменателем, равным единице, и, по определению произведения дробей, получим

$$3 \cdot \frac{5}{7} = \frac{3}{1} \cdot \frac{5}{7} = \frac{3 \cdot 5}{1 \cdot 7} = \frac{15}{7}.$$

2. Разделить целое число 2 на рациональное число 1/3. Записывая целое число 2 в виде дроби со знаменателем, равным единице, и пользуясь определением частного двух дробей, получаем

$$2 : \frac{1}{3} = \frac{2}{1} : \frac{1}{3} = \frac{2 \cdot 3}{1 \cdot 1} = \frac{6}{1} = 6.$$

Свойство ассоциативности умножения рациональных чисел позволяет ввести понятие натуральной степени рационального числа: *k-й степенью* рационального числа m/n называется рациональное число q, полученное в результате умножения числа самого на себя k раз:

$$q = \underbrace{\frac{m}{n} \cdot \frac{m}{n} \cdot \ldots \cdot \frac{m}{n}}_{k \text{ сомножителей}}.$$

Для обозначения k-й степени числа m/n обычно используется запись

$$q = \left(\frac{m}{n}\right)^k,$$

в которой число m/n называется *основанием* степени, а число k — *показателем* степени.

3.3. Множество рациональных чисел как множество упорядоченных пар целых чисел. Множество всех целых чисел Z замкнуто относительно операций сложения, умножения и вычитания, но не замкнуто относительно операции деления: сумма, разность и произведение двух целых чисел есть число целое, в то время как частное m/n двух целых чисел m и n будет целым числом лишь в том случае, если число $|m|$ кратно числу n. Множество рациональных чисел получается как расширение множества целых чисел добавлением новых числовых объектов, таких, чтобы:

1) расширенное множество содержало в качестве своего собственного подмножества множество всех целых чисел;

2) арифметические операции, определенные для целых чисел, были определены и для элементов расширенного множества, причем результат этих операций для целых чисел, рассматриваемых как элементы расширенного множества, должен совпадать с тем, какой они имели в множестве целых чисел;

3) в расширенном множестве была выполнима операция деления (кроме деления на нуль), которая в множестве целых чисел не всегда выполнима, — частное двух элементов расширенного множества должно являться элементом этого множества;

4) расширенное множество было минимальным в том смысле, что оно не должно содержать собственного подмножества, удовлетворяющего условиям 1)—3).

Существует единственное множество, удовлетворяющее условиям 1)—4). Этим множеством является упорядоченное множество всех рациональных чисел с введенными в нем арифметическими операциями.

В качестве метода пополнения множества всех целых чисел до множества всех рациональных чисел может быть выбран способ представления рациональ-

ных чисел как упорядоченных пар целых чисел (т. е. как рациональных дробей с определенными правилами сравнения и правилами арифметических действий (см. п. 3.1)). При таком определении рациональных чисел множество всех целых чисел есть подмножество множества всех рациональных чисел, если считать, что любая пара $(m; n)$ при $m=k\cdot n$, где k – целое число, отождествляется с числом k. В частности, все нулевые пары $(0; n)$ отождествляются с числом нуль.

При таком способе отождествления пары $(m; n)$ с целым числом результаты операций сложения, вычитания и умножения пар $(m; n)$ и $(p; q)$ при $m=k\cdot n$ и $p=r\ q$ совпадают с результатами операций сложения, вычитания и умножения целых чисел k и r.

Действительно, пусть $(m; n)$ и $(p; q)$ – две пары такие, что число $|m|$ кратно n, а число $|p|$ кратно q, т. е.

$$m=k\cdot n \quad \text{и} \quad p=l\ q,$$

где k и l – отличные от нуля целые числа. По определению суммы пар имеем

$$(m; n)+(p; q)\sim(m\cdot q+n\cdot p; n\cdot q).$$

Но $m=k\ n$, а $p=l\cdot q$, поэтому пара $(m\cdot q+n\cdot p; n\cdot q)$ отождествляется с целым числом $k+l$.

С другой стороны, пара $(m; n)$ отождествляется с целым числом k, а пара $(p; q)$ – с целым числом l; следовательно, сумма этих пар, как сумма целых чисел, также равна $k+l$.

Аналогично проверяется и совпадение результатов умножения и вычитания пар с соответствующими результатами умножения и вычитания целых чисел.

Из определения операции деления пар следует, что частное двух целых чисел (из которых второе отлично от нуля), рассматриваемых как упорядоченные пары (т. е. как элементы расширенного множества), всегда существует в множестве пар, т. е. пара $(m; n)$, полученная как частное двух целых чисел, существует для любых целых чисел m, n $(n\neq 0)$, а не только для $m=k\ n$, где k – целое. Действительно, используя определение операций умножения и деления пар, для любых целых чисел m и n $(n\neq 0)$, записанных в виде пар $(m; 1)$ и $(n; 1)$, имеем

$$(m; 1):(n; 1)\sim(m; 1)\cdot(1; n)\sim(m; n).$$

Таким образом, в тех случаях, когда m не делится на n, пара $(m; n)$ представляет собой не совпадающее ни с каким целым числом рациональное число, являющееся частным от деления целого числа m на целое число n.

§ 4. Десятичные дроби

4.1. Десятичная позиционная система счисления. Наиболее распространенный способ записи чисел — это запись чисел в *десятичной системе счисления* *). Суть этого способа записи состоит в следующем.

1. Все натуральные числа от единицы до девяти обозначаются следующими индивидуальными символами:

$$1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9.$$

К этим символам дополнительно присоединяется десятый символ 0, называемый *нулем*. Перечисленные десять символов называются *цифрами* десятичной системы счисления.

2. Одна и та же цифра имеет разный смысл в зависимости от расположения этой цифры относительно других цифр, участвующих в записи числа (в этом и заключается позиционность системы счис-

*) Обычно слово «позиционная» опускают и говорят просто «десятичная система счисления».

ления). Так, в позиционной десятичной системе счисления комбинации 4521 и 4125 четырех различных цифр 1, 2, 4, 5 дают запись двух различных натуральных чисел.

Цифра, стоящая на первом месте справа в записи натурального числа, указывает на количество единиц, содержащихся в данном числе, на втором — количество десятков, на третьем — количество сотен, на четвертом — количество тысяч и т. д.

При записи натурального числа в десятичной позиционной системе счисления о первой цифре, стоящей справа, говорят, что она стоит в *разряде единиц*, о второй — что она стоит в *разряде десятков*, о третьей — что она стоит в *разряде сотен*, и т. д.

В *десятичной системе счисления* запись

$$n_k n_{k-1} n_{k-2} \ldots n_1 n_0,$$

где n_k, n_{k-1}, n_{k-2}, ..., n_1, n_0 — цифры, есть условная запись числа

$$n_k \cdot 10^k + n_{k-1} \cdot 10^{k-1} + n_{k-2} \cdot 10^{k-2} + \ldots + n_1 \cdot 10 + n_0. \qquad (1)$$

При записи натурального числа в десятичной системе счисления также принято следующее соглашение: крайней левой цифрой может быть любая из десяти цифр, кроме цифры нуль. Так, принято записывать число двенадцать в виде 12, а не в виде 012 или 0012. С помощью этого соглашения избавляются от возможной неоднозначности в записи одного и того же числа.

4.2. Понятие десятичной дроби. *Десятичная дробь* — это форма записи действительного числа в десятичной позиционной системе счисления. Десятичная дробь

$$N, n_1 n_2 n_3 \ldots n_k \ldots, \qquad (2)$$

где N — целое число, а n_1, n_2, n_3, ..., n_k, ... — цифры десятичной системы счисления, — это условная запись действительного числа

$$N + \frac{n_1}{10^1} + \frac{n_2}{10^2} + \frac{n_3}{10^3} + \ldots + \frac{n_k}{10^k} + \ldots$$

Целое число N называется *целой частью* десятичной дроби (2). О цифре n_1, стоящей в десятичной дроби (2) на первом месте после запятой, говорят, что она стоит в *разряде десятых долей* единицы; о цифре n_2, стоящей на втором месте, говорят, что она стоит в *разряде сотых долей* единицы; о цифре n_3 — что она стоит в *разряде тысячных долей*, и т. д.

В соответствии с данным выше определением десятичной дроби, в случае, если N — отрицательное целое число, знак минус ставят не перед всей дробью, а над числом $|N|$. Например, десятичная дробь $\bar{2},135\ldots$ есть условная запись действительного числа

$$-2 + 0,135\ldots$$

Следует отметить, что при отрицательном N такая запись десятичной дроби в ряде случаев (например, при выполнении арифметических действий) весьма неудобна. Поэтому часто используют другую форму записи.

Пусть a — положительное действительное число, имеющее десятичное представление

$$a = 3,521.$$

Число $-a$, противоположное числу a, в соответствии с принятым выше обозначением будет иметь десятичное представление

$$-a=\overline{4},479.$$

Чтобы избежать несоответствия в обозначениях действительных чисел и их десятичных представлений, число $-a$ записывают также в виде

$$-a=-3,521=-3-\frac{5}{10^1}-\frac{2}{10^2}-\frac{1}{10^3}.$$

Десятичная дробь $N, n_1n_2n_3\ldots n_k\ldots$ называется *бесконечной*, если для любого натурального k найдется такой номер $l>k$, что $n_l\neq 0$.

Десятичная дробь $N, n_1n_2n_3\ldots n_k\ldots$ называется *конечной*, если найдется такое натуральное число k, что $n_k\neq 0$ и $n_l=0$ при всех $l>k$. Цифры числа N и $n_1, n_2, \ldots, n_k$ называют *значащими цифрами*.

Обычно в записи конечной дроби нули, стоящие после последней значащей цифры, опускают, т. е. дробь

$$N,n_1n_2n_3\ldots n_k00\ldots0\ldots,$$

где $n_k\neq 0$, записывают в виде

$$N,n_1n_2n_3\ldots n_k.$$

Периодические десятичные дроби. Бесконечная десятичная дробь $N,n_1n_2n_3\ldots n_k\ldots$ называется *периодической*, если существуют такие натуральные числа p, q, что $n_{k+p}=n_k$ для всех $k>q$. Для обозначения бесконечной десятичной периодической дроби используется запись $N,n_1n_2\ldots n_q(n_{q+1}n_{q+2}\ldots n_{q+p})$, где совокупность цифр $n_{q+1}n_{q+2}\ldots n_{q+p}$ называется *периодом дроби*.

Иногда периодические дроби разделяют на *чисто периодические*, которые записываются в виде

$$N,(n_1n_2\ldots n_p),$$

и *смешанные периодические*, которые записываются в виде

$$N,n_1n_2\ldots n_k(n_{k+1}n_{k+2}\ldots n_{k+p}) \qquad (k\geqslant 1).$$

Например, дробь $2,131313\ldots=2,(13)$ — чисто периодическая, а дробь $2,41313\ldots=2,4(13)$ — смешанная периодическая.

Разбиение множества десятичных дробей на подмножества конечных и бесконечных десятичных дробей весьма условно, так как любая конечная десятичная дробь (исключая число нуль) может быть записана как бесконечная периодическая дробь. Это можно сделать совсем очевидным способом (представить, например, конечную десятичную дробь 0,25 как $0,25000\ldots0\ldots=0,25(0)$, т. е. присоединить к конечной десятичной дроби бесконечно много нулей).

Бесконечные десятичные дроби, кроме тех, период которых состоит из одной цифры 9, называются *допустимыми десятичными дробями*.

В множестве допустимых десятичных дробей может быть введено *отношение порядка*, т. е. множество всех допустимых десятичных дробей является *упорядоченным множеством*.

Две допустимые бесконечные дроби

$$N,n_1n_2n_3\ldots n_k\ldots \quad \text{и} \quad M,m_1m_2m_3\ldots m_k\ldots$$

считаются *равными*, если целая часть первой дроби равна целой части второй дроби ($N = M$) и цифры, стоящие на одних и тех же местах в записях этих дробей, одинаковы:

$$n_1 = m_1, \quad n_2 = m_2, \quad \ldots \quad n_k = m_k, \quad \ldots$$

Допустимая бесконечная десятичная дробь $0, n_1n_2n_3\ldots n_k\ldots$ считается *меньше* допустимой бесконечной дроби $0, m_1m_2m_3\ldots m_k\ldots$:

$$0, n_1n_2n_3\ldots n_k\ldots < 0,m_1m_2m_3\ldots m_k\ldots$$

если найдется такой номер k, что

$$n_1 = m_1, \quad n_2 = m_2, \quad \ldots, \quad n_{k-1} = m_{k-1}, \quad \text{но} \quad n_k < m_k,$$

т. е. сравнение допустимых десятичных дробей проводится по первой встретившейся паре неравных цифр. Например, дробь 0,72159... меньше дроби 0,72160..., так как первые три цифры, стоящие после запятой, у обеих дробей совпадают, а четвертая цифра первой дроби меньше четвертой цифры второй дроби.

Допустимая десятичная дробь $N,n_1n_2n_3\ldots n_k\ldots$ считается меньше допустимой десятичной дроби $M,m_1m_2m_3\ldots m_k\ldots$, если целая часть первой дроби меньше целой части второй дроби ($N < M$), либо, если целые числа этих двух дробей равны ($N = M$), десятичная дробь $0, n_1n_2n_3\ldots n_k\ldots$ меньше десятичной дроби $0,m_1m_2m_3\ldots m_k\ldots$

Сравнение двух конечных, а также конечной и допустимой бесконечной десятичных дробей производится по тем же правилам, что и сравнение двух допустимых бесконечных десятичных дробей.

4.3. Арифметические действия с конечными десятичными дробями. Арифметические действия с конечными десятичными дробями будем производить, считая, что если положительное число a может быть записано в виде конечной десятичной дроби

$$N_0,n_1n_2\ldots n_k,$$

то отрицательное число $-a$ записывается в виде конечной десятичной дроби

$$-N_0,n_1n_2\ldots n_k.$$

Сложение и вычитание конечных десятичных дробей выполняется так же, как сложение и вычитание целых чисел; необходимо только записывать каждый разряд одной дроби под разрядом того же наименования второй дроби, а на место пустующих разрядов ставить нули.

Пример 1. Сложить дроби 2,14 и 0,151.

Запишем дробь 2,14 в виде 2,140 и произведем сложение целых чисел 2140 и 151:

$$2140 + 151 = 2291.$$

Отделив справа у целого числа 2291 то же самое число знаков, что было отделено у дробей 2,140 и 0,151 (а именно три знака), получим дробь 2,291, которая является суммой данных дробей.

Умножение конечных десятичных дробей производится следующим образом: не обращая внимания на запятые, конечные дроби перемножают, как целые числа; в получившемся произведении отделяют справа число знаков, равное сумме числа знаков после запятой у всех сомножителей.

Пример 2. Найти произведение двух конечных десятичных дробей 2,1 и 0,27.

Перемножив целые числа 21 и 27 (нуль в числе 027 можно отбросить), получим число 567. Отделив справа три знака, получим произведение двух данных дробей, равное 0,567.

Деление конечной десятичной дроби на целое число производится следующим образом.

1) Если делимое меньше делителя, пишем в целой части частного нуль и ставим после него запятую. Затем, не обращая внимания на запятую в делимом, присоединяем к целой части делимого первую цифру его дробной части; если после такого присоединения получается число, меньшее делителя, ставим после запятой в частном нуль и присоединяем к полученному ранее числу следующую цифру делимого; если и после этого получаем число, меньшее делителя, ставим еще нуль и т. д., пока не получим число, превосходящее делитель. В дальнейшем деление совершается так же, как с целыми числами, причем делимое можно неограниченно «расширять» вправо от запятой, приписывая в конце нули.

Пример 3. Разделить десятичную дробь 0,525 на число 2.

В согласии с приведенным правилом вычисление частного будет выглядеть так:

$$\begin{array}{r|l} 0{,}525 & 2 \\ \hline -5 & 0{,}2625 \\ \quad 4 & \\ \hline \quad 12 & \\ -12 & \\ \hline \quad 5 & \\ \quad -4 & \\ \hline \quad 10 & \\ \quad -10 & \\ \hline \quad 0 & \end{array}$$

Может оказаться, что описанный процесс деления никогда не закончится. В таком случае частное нельзя выразить конечной десятичной дробью.

2) Если делимое больше делителя, делим сначала целую часть, записываем в частном результат деления и ставим запятую. После этого деление продолжается, как в предыдущем случае.

Деление конечной десятичной дроби на конечную десятичную дробь производится по следующему правилу.

Чтобы разделить десятичную дробь (или целое число) на десятичную дробь, опускаем запятую в делителе; в делимом же переносим запятую вправо на столько знаков, сколько их было в дробной части делителя (в случае необходимости к делимому в конце приписываем нули). После этого выполняем деление десятичной дроби на целое число, как указано выше. Так, деление десятичной дроби 0,525 на дробь 0,2 сводится к делению десятичной дроби 5,25 на число 2.

Арифметические действия над бесконечными периодическими десятичными дробями довольно сложны и громоздки, а потому гораздо проще поступать следующим образом: перевести бесконечные периодические десятичные дроби, над которыми требуется произвести арифметические действия, в рациональные дроби; над рациональными дробями произвести необходимые операции; дробь, получившуюся в результате вычислений, если это необходимо, перевести в десятичную.

4.4. Обращение конечной десятичной дроби в рациональную дробь. Пусть $N,n_1n_2\ldots n_k$ — конечная десятичная дробь. Десятичная дробь

$0,n_1n_2\ldots n_k$ может быть записана в виде

$$\frac{n_1n_2\ldots n_k}{10^k}.$$

К этой дроби прибавим целую часть данной десятичной дроби. В результате получим рациональную дробь.

Например, конечную десятичную дробь 2,135 можно обратить в рациональную:

$$2{,}135 = 2 + 0{,}135 = 2 + \frac{135}{10^3} = \frac{2135}{1000}.$$

4.5. Обращение бесконечной периодической дроби в рациональную дробь.

1) Представляем бесконечную периодическую дробь $0,n_1n_2\ldots n_k(n_{k+1}\ldots n_{k+p})$ в виде суммы конечной десятичной дроби и бесконечной периодической дроби:

$$0,n_1n_2\ldots n_k(n_{k+1}\ldots n_{k+p}) = 0,n_1n_2\ldots n_k + 0,\underbrace{0\ldots0}_{k}(n_{k+1}\ldots n_{k+p}).$$

2) Бесконечную периодическую дробь $0,\underbrace{0\ldots0}_{k}(n_{k+1}\ldots n_{k+p})$ представляем в виде произведения:

$$0,\underbrace{00\ldots0}_{k}(n_{k+1}\ldots n_{k+p}) = \frac{1}{10^k}\cdot 0,(n_{k+1}n_{k+2}\ldots n_{k+p}). \qquad (3)$$

3) Чисто периодическую дробь $0,(n_{k+1}n_{k+2}\ldots n_{k+p})$ записываем в виде

$$0,(n_{k+1}n_{k+2}\ldots n_{k+p}) = \frac{n_{k+1}n_{k+2}\ldots n_{k+p}}{10^p} + \frac{n_{k+1}n_{k+2}\ldots n_{k+p}}{10^{2p}} + \ldots$$
$$\ldots + \frac{n_{k+1}n_{k+2}\ldots n_{k+p}}{10^p}\cdot\left(1 + \frac{1}{10^p} + \frac{1}{10^{2p}} + \ldots\right).$$

4) Вычисляем сумму членов бесконечно убывающей геометрической прогрессии, стоящей в квадратных скобках. Она равна

$$\frac{10^p}{10^p-1},$$

и, следовательно, чисто периодическая дробь $0,(n_{k+1}n_{k+2}\ldots n_{k+p})$ обращается в следующую рациональную дробь:

$$0,(n_{k+1}n_{k+2}\ldots n_{k+p}) = \frac{n_{k+1}n_{k+2}\ldots n_{k+p}}{10^p}\cdot\frac{10^p}{10^p-1} = \frac{n_{k+1}n_{k+2}\ldots n_{k+p}}{10^p-1}.$$

Теперь с помощью равенства (3) нетрудно обратить десятичную дробь $0,n_1n_2\ldots n_k(n_{k+1}n_{k+2}\ldots n_{k+p})$ в рациональную дробь:

$$0,n_1n_2\ldots n_k(n_{k+1}n_{k+2}\ldots n_{k+p}) =$$
$$= 0,n_1n_2\ldots n_k + \frac{1}{10^k}\cdot\frac{n_{k+1}n_{k+2}\ldots n_{k+p}}{10^p-1} =$$
$$= \frac{n_1n_2\ldots n_kn_{k+1}n_{k+2}\ldots n_{k+p} - n_1n_2\ldots n_k}{10^k(10^p-1)}.$$

Правило перевода смешанной периодической десятичной дроби в рациональную дробь можно сформулировать так.

Чтобы обратить смешанную периодическую десятичную дробь в рациональную, нужно из числа, образованного цифрами, стоящими после запятой до начала второго периода, вычесть число, образованное из цифр, стоящих после запятой до начала первого периода; полученную разность взять в качестве числителя дроби, а в знаменателе написать цифру 9 столько раз, сколько цифр в периоде, и со столькими нулями, сколько цифр между запятой и началом первого периода.

Примеры. 1. Перевести чисто периодическую дробь 2,(13) в рациональную дробь.

Представим дробь в виде

$$2,(13)=2+0,(13)=2+\frac{13}{100}+\frac{13}{10\,000}+\ldots=2+\frac{13}{10^2}+\frac{13}{10^4}+\ldots=$$
$$=2+\frac{13}{10^2}\left(1+\frac{1}{10^2}+\frac{1}{10^4}+\ldots\right).$$

В скобках стоит сумма членов бесконечно убывающей геометрической прогрессии с первым членом, равным 1, и знаменателем $q=1/10^2$. Используя формулу для суммы членов бесконечно убывающей геометрической прогрессии, получаем

$$2,(13)+\frac{13}{10^2}\cdot\frac{1}{1-1/10^2}=2+\frac{13}{10^2}\cdot\frac{10^2}{10^2-1}=2+\frac{13}{99}=\frac{211}{99}.$$

Таким образом, периодическая десятичная дробь 2,(13) может быть обращена в рациональную дробь 211/99.

2. Перевести смешанную периодическую дробь 2,5(13) в рациональную дробь.

Представим дробь 2,5(13) в виде

$$2,5(13)=2+0,5+0,0(13)=2+0,5+\frac{1}{10}\cdot 0,(13).$$

Далее необходимо перевести чисто периодическую дробь 0,(13) в рациональную дробь аналогично тому, как это было сделано в примере 1, и произвести сложение.

4.6. Десятичные представления рациональных чисел. Рациональное число определяется как число, которое может быть записано в виде m/n, где m и n — целые числа и $n>0$. Рассмотрим три примера представления различных рациональных чисел в виде десятичных дробей.

Пример 1. Записать рациональное число 1/4 в виде десятичной дроби:

$$\begin{array}{r|l} 1 & 4 \\ \hline 10 & 0,25 \\ -\,8 & \\ \hline 20 & \\ -\,20 & \\ \hline 0 & \end{array}$$

Из проведенных вычислений видно, что процедура деления заканчивается через конечное число шагов и число 1/4 может быть записано в виде конечной десятичной дроби 0,25.

Пример 2. Записать рациональное число 5/11 в виде десятичной дроби:

```
     5      | 11
  ---------  |-----
*  50       | 0,4545
  -44
  ---
    60
   -55
  ---
*   50
   -44
  ---
    60
   -55
  ---
*    5
```

Из проведенных вычислений следует, что если продолжать процесс деления, то и дальше в остатке будут последовательно получаться числа 6 и 5, и каждый раз, при появлении в остатке числа 5, цикл деления будет начинаться заново (в разобранном примере звездочками отмечены начала первых трех таких циклов). В этом примере процедура деления никогда не закончится, следовательно, рациональное число 5/11 записывается в виде бесконечной десятичной периодической дроби 0,(45).

Пример 3. Записать рациональное число 131/990 в виде десятичной дроби:

```
     131   | 990
  ---------  |-----
  1310      | 0,132
  -990
  ---
*  3200
  -2970
  ---
    2300
   -1980
  ---
*    320
```

В этом примере, так же как и в предыдущем, процедура деления никогда не закончится. Однако здесь повторение цикла деления начинается не с того числа, с которого начиналось деление, а с другого числа (здесь числа 320), которое появилось в одном из остатков. Следовательно, рациональное число 131/990 записывается в виде смешанной периодической десятичной дроби 0,1(32).

Три разобранных примера описывают все возможные случаи, которые встречаются при записи рационального числа в виде десятичной дроби. В самом деле, рассмотрим произвольное положительное число m/n, где m и n — натуральные числа (случай отрицательного рационального числа может быть разобран аналогично). При делении натурального числа m на натуральное число n в остатке могут появиться лишь следующие числа:

$$0,\ 1,\ 2,\ \ldots,\ n-1.$$

Если в процессе деления в остатке появится число 0, то процесс деления будет закончен и, следовательно, рациональное число m/n будет представлено в виде конечной десятичной дроби. Если же в процессе деления число 0 в остатке не появится, то по крайней мере через $n-1$ шаг неизбежно возникнет повторение остатка; с этого места начинается новый цикл. Результатом деления является беско-

нечная периодическая десятичная дробь (точнее говоря, допустимая периодическая десятичная дробь).

Таким образом, всякое рациональное число m/n представимо либо в виде конечной, либо в виде бесконечной периодической дроби; и обратно, любая конечная, а также любая бесконечная периодическая десятичная дробь есть запись некоторого рационального числа.

Рациональное число m/n (m, n — целые, взаимно простые числа и $n > 1$) может быть записано в виде конечной десятичной дроби тогда и только тогда, когда число n имеет своими простыми делителями лишь числа 2 и 5. При этом число n не обязано иметь среди своих простых делителей как число 2, так и число 5; оно может делиться лишь на одно из них. Если $n = 1$, то дробь m/n также, очевидно, записывается в виде конечной десятичной дроби. Например, рациональные числа 1/25, 1/16 и 7/1, где n равно соответственно 25, 16 и 1, представляются в виде конечных десятичных дробей:

$$\frac{1}{25}=0{,}04; \quad \frac{1}{16}=0{,}0625; \quad \frac{7}{1}=7.$$

4.7. Непрерывные дроби. Алгоритм Евклида (см. п. 1.5) нахождения наибольшего общего делителя двух натуральных чисел приводит к записи рациональных чисел в виде *непрерывных дробей*.

Например, применение алгоритма Евклида к числам 840 и 611 дает следующий ряд равенств:

$$\begin{aligned} 840&=1\cdot 611+229,\\ 611&=2\cdot 229+153,\\ 229&=1\cdot 153+76,\\ 153&=2\cdot 76+1, \end{aligned}$$

которые можно записать в виде

$$\frac{840}{611}=1+\frac{229}{611}=1+\cfrac{1}{\cfrac{611}{229}}, \quad \frac{229}{153}=1+\frac{76}{153}=1+\cfrac{1}{\cfrac{153}{76}},$$

$$\frac{611}{229}=2+\frac{153}{229}=2+\cfrac{1}{\cfrac{229}{153}}, \quad \frac{153}{76}=2+\frac{1}{76}.$$

Комбинируя последние равенства, приходим к следующему представлению рационального числа 840/611:

$$\frac{840}{611}=1+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{76}}}}.$$

Рассмотрим теперь цепочку равенств алгоритма Евклида нахождения наибольшего общего делителя двух натуральных чисел m и n:

$$\begin{aligned} m&=n\cdot m_1+r_1,\\ n&=r_1\cdot m_2+r_2,\\ r_1&=r_2\,m_3+r_3,\\ &\ldots\ldots\ldots\\ r_{k-2}&=r_{k-1}\cdot m_k+r_k,\\ r_{k-1}&=r_k\,m_{k+1}. \end{aligned}$$

где $n > r_1 > r_2 > r_3 > \ldots > r_k > 0$.

Перепишем эту цепочку равенств в виде

$$\frac{m}{n}=m_1+\frac{r_1}{n},$$

$$\frac{n}{r_1}=m_2+\frac{r_2}{r_1},$$

$$\frac{r_1}{r_2}=m_3+\frac{r_3}{r_2},$$

$$\cdots\cdots\cdots\cdots$$

$$\frac{r_{k-2}}{r_{k-1}}=m_k+\frac{r_k}{r_{k-1}},$$

$$\frac{r_{k-1}}{r_k}=m_{k+1}.$$

Каждое из приведенных равенств (за исключением последнего) описывает выделение целой части неправильной дроби, т. е. представляет неправильную дробь в виде суммы натурального числа и некоторой правильной дроби. Замечая, что левая часть каждого равенства есть величина, обратная второму слагаемому правой части предыдущего равенства, рациональное число m/n можно записать, используя лишь числа $m_1, m_2, m_3, \ldots, m_k, m_{k+1}$:

$$\frac{m}{n}=m_1+\cfrac{1}{m_2+\cfrac{1}{m_3+\cfrac{1}{m_4+\ddots+\cfrac{1}{m_{k+1}}}}}, \qquad (4)$$

где $m_2, m_3, m_4, \ldots, m_{k+1}$ — натуральные числа, а m_1 — либо натуральное число, либо нуль.

Выражение, стоящее в правой части равенства (4), называется *конечной непрерывной* (или *цепной*) *дробью*.

Если рациональное число m/n отрицательно ($m<0$, $n>0$), то при записи числа m/n в виде непрерывной дроби используют следующий прием. Отрицательное рациональное число m/n представляют в виде

$$\frac{m}{n}=M+\frac{k}{n},$$

где M — отрицательное число, а k — натуральное число, меньшее n. Положительную рациональную дробь k/n записывают в виде непрерывной дроби:

$$\frac{k}{n}=\cfrac{1}{l_1+\cfrac{1}{l_2+\cfrac{1}{l_3+\ddots+\cfrac{1}{l_k}}}}.$$

Тогда отрицательная рациональная дробь m/n записывается в виде

$$\frac{m}{n}=M+\cfrac{1}{l_1+\cfrac{1}{l_2+\cfrac{1}{l_3+\ddots+\cfrac{1}{l_k}}}}.$$

Непрерывную дробь, стоящую в правой части равенства (4), обозначают

$$[m_1;\ m_2,\ m_3,\ \ldots,\ m_{k+1}], \tag{5}$$

где точка с запятой отделяет целую часть рационального числа от его дробной части.

Непрерывная дробь

$$[m_1;\ m_2,\ m_3,\ \ldots,\ m_s], \quad s \leqslant k+1,$$

называется *подходящей дробью* s-го порядка для непрерывной дроби (5) и кратко обозначается p_s/q_s.

Подходящая непрерывная дробь называется *четной*, если s четно, и *нечетной*, если s нечетно. Все подходящие дроби четного порядка меньше m/n, и величина их возрастает с ростом порядка; все подходящие дроби нечетного порядка больше m/n, и величина их убывает с ростом порядка;

$$\frac{p_s}{q_s} < \frac{m}{n} < \frac{p_{s+1}}{q_{s+1}}, \quad s \text{—четное}.$$

Модуль разности между подходящими дробями p_s/q_s и p_{s+1}/q_{s+1} равен $\dfrac{1}{q_s \cdot q_{s+1}}$, а потому разность между подходящей дробью s-го порядка и числом m/n можно оценить неравенством

$$\left| \frac{m}{n} - \frac{p_s}{q_s} \right| < \frac{1}{q_s \cdot q_{s+1}}. \tag{6}$$

Теория непрерывных дробей исторически возникла из потребности замены рациональной дроби с большими числителем и знаменателем другой рациональной дробью, у которой числитель и знаменатель были бы значительно меньше (и меньше некоторого наперед заданного числа) и такой, чтобы она по своей величине как можно меньше отличалась от исходной дроби. Пусть m и n — числа, равные, например, 1355 и 946 соответственно. Необходимо заменить дробь 1355/946 дробью p_k/q_k, где q_k должно быть меньше 100, а также оценить погрешность, возникающую при такой замене. Представим дробь 1355/946 в виде непрерывной дроби и вычислим подходящие дроби. При этом окажется, что

$$\frac{m}{n} = \frac{1355}{946} = [1;\ 2,\ 3,\ 5,\ 8,\ 3], \quad \frac{p_3}{q_3} = \frac{53}{37}, \quad \frac{p_4}{q_4} = \frac{434}{303}.$$

Нашему условию (знаменатель меньше 100) удовлетворяет дробь 57/37.

Неравенство (6) позволяет просто оценить абсолютную погрешность между данной дробью и подходящей дробью третьего порядка:

$$\left| \frac{1355}{946} - \frac{53}{37} \right| < \frac{1}{q_3 \cdot q_4} = \frac{1}{11211} < 0{,}0001.$$

В теории непрерывных дробей вводится также понятие бесконечной непрерывной дроби и доказывается, что любое действительное число может быть единственным образом записано в виде непрерывной дроби. При этом рациональные числа записываются в виде конечных, а иррациональные — в виде бесконечных непрерывных дробей.

Заметим, что как в случае рационального, так и в случае иррационального числа подходящие непрерывные дроби обладают тем свойством, что они дают «наилучшее» приближение любого действительного числа рациональным числом. А именно: для любой подходящей дроби k-го порядка p_k/q_k нельзя подобрать рациональную дробь со знаменателем, меньшим чем q_k, такую, чтобы она давала лучшее приближение действительного числа a, чем подходящая дробь p_k/q_k.

§ 5. Действительные числа

5.1. Множество действительных чисел как расширение множества рациональных чисел. В § 2, 3 было последовательно проведено расширение множества натуральных чисел до множества целых чисел и расширение множества целых чисел до множества рациональных чисел. Множество всех рациональных чисел представляет собой множество, замкнутое относительно операций сложения, умножения, вычитания и деления (кроме деления на нуль),— сумма, произведение, разность и частное двух рациональных чисел опять будет рациональным числом.

Однако оказывается, что существуют алгебраические и геометрические задачи, которые не имеют решения в множестве рациональных чисел. Так, задачей, часто не имеющей решения в множестве рациональных чисел, является извлечение корня из положительного целого числа. Например, число $\sqrt{2}$ (см. п. 5.9) не является рациональным числом, т. е. его нельзя записать в виде m/n, где m и n — целые числа и $n \neq 0$. Нетрудно привести и другие примеры чисел, которые не могут быть представлены в виде m/n, т. е. не являются рациональными числами.

Множество всех действительных чисел также может быть получено как естественное расширение множества всех рациональных чисел. Однако в отличие от довольно простых способов расширений множества натуральных чисел до множества целых чисел и множества целых чисел до множества рациональных чисел, метод расширения (или пополнения) множества всех рациональных чисел до множества действительных чисел оказывается более сложным и основан на использовании понятий математического анализа. Математически строгая теория действительных чисел была развита лишь в середине прошлого века в трудах Р. Дедекинда и Г. Кантора, и при ее создании был использован ряд весьма тонких результатов математического анализа.

Один из методов пополнения множества рациональных чисел до множества действительных чисел основан на понятии фундаментальной последовательности рациональных чисел. Покажем схематически, в чем состоит суть этого метода.

Последовательность (x_n) рациональных чисел x_n называется *фундаментальной*, если для любого рационального $\varepsilon > 0$ существует такое натуральное число n_0, что

$$|x_p - x_q| < \varepsilon$$

для всех p и q, больших n_0. Используя определение сходящейся последовательности, нетрудно доказать, что всякая сходящаяся последовательность будет фундаментальной. С другой стороны, всякая фундаментальная последовательность рациональных чисел будет иметь предел, однако может оказаться, что этот предел не будет рациональным числом. Примером такой последовательности может служить последовательность положительных рациональных чисел (r_n), квадраты которых становятся сколь угодно близкими числу 2:

$$|r_1^2 - 2| < \frac{1}{10}; \quad |r_2^2 - 2| < \frac{1}{10^2}; \ \ldots; \ |r_n^2 - 2| < \frac{1}{10^n}; \ \ldots$$

Из способа задания последовательности (r_n) следует, что ее предел — число, не принадлежащее множеству рациональных чисел; это есть иррациональное число $\sqrt{2}$.

Множество всех действительных чисел получается пополнением множества рациональных чисел новыми числовыми объектами — *иррациональными числами*,

которые являются пределами всевозможных фундаментальных последовательностей рациональных чисел. Множество действительных чисел обычно обозначается буквой R.

5.2. Аксиоматическое построение множества действительных чисел. Множество всех действительных чисел может быть описано как множество, элементы которого удовлетворяют перечисленным ниже свойствам I—VI. Элементы этого множества будем называть числами.

I. Свойство упорядоченности Для любых двух чисел a и b определено *отношение порядка*, т. е. для любых двух действительных чисел a и b

либо $a=b$ (a равно b),

либо $a<b$ (a меньше b),

либо $b<a$ (b меньше a);

причем если $a<b$ и $b<c$, то $a<c$.

II. Свойства операции сложения В множестве действительных чисел определена бинарная операция *сложения*, т. е. любой упорядоченной паре чисел $(a; b)$ ставится в соответствие число, обозначаемое $a+b$ и называемое *суммой* чисел a и b.

При этом:

1) для любой тройки чисел a, b, c

$$(a+b)+c=a+(b+c) \quad (\textit{ассоциативный закон сложения});$$

2) для любой пары чисел a и b

$$a+b=b+a \quad (\textit{коммутативный закон сложения});$$

3) существует число, обозначаемое символом 0 и называемое *нулем*, такое, что для любого числа a

$$a+0=a;$$

4) для любого числа a существует число, обозначаемое $-a$, такое, что

$$a+(-a)=0;$$

число $-a$ называется *противоположным* числу a;

5) если $a<b$, то для любого числа c

$$a+c<b+c.$$

Число $a>0$ называется *положительным*, а число $a<0$ — *отрицательным*.

Для любой упорядоченной пары чисел $(a; b)$ число $a+(-b)$ называется *разностью* чисел a и b и обозначается $a-b$.

III. Свойства операции умножения В множестве действительных чисел определена бинарная операция, называемая *умножением*, т. е. любой упорядоченной паре чисел $(a; b)$ ставится в соответствие число, обозначаемое $a\cdot b$ и называемое *произведением* чисел a и b.

При этом:

1) для любой тройки чисел a, b, c

$$(a\cdot b)\cdot c=a\cdot(b\cdot c) \quad (\textit{ассоциативный закон умножения});$$

2) для любой пары чисел a, b

$$a\cdot b=b\cdot a \quad (\textit{коммутативный закон умножения});$$

3) существует число, обозначаемое символом 1 и называемое *единицей*, такое, что для любого числа a

$$a \cdot 1 = a;$$

4) для любого числа a, отличного от нуля, существует число, обозначаемое $1/a$, такое, что

$$a \cdot \frac{1}{a} = 1;$$

число $1/a$ называется *обратным* числу a;

5) если $a < b$ и $c > 0$, то $a \cdot c < b \cdot c$; если $a < b$ и $c < 0$, то

$$a \cdot c > b \cdot c.$$

Для любой упорядоченной пары чисел a и b (b отлично от нуля) число $a \cdot \frac{1}{b}$ называется *частным* от деления a на b и обозначается $\frac{a}{b}$:

$$\frac{a}{b} = a \cdot \frac{1}{b}.$$

IV. Связь операций сложения и умножения. Для любой тройки чисел a, b и c

$$(a + b) \cdot c = a \cdot c + b \cdot c$$

(*дистрибутивность умножения относительно сложения*).

V. Свойство Архимеда. Для любого числа a существует такое целое число n, что $n > a$.

Из этого свойства, в частности, следует, что для любых чисел a и b при $a > 0$ существует такое натуральное число n, что

$$n \cdot a > b.$$

VI. Свойство непрерывности. Перечисленные выше свойства I—V были присущи и некоторым другим числовым множествам (например, множеству рациональных чисел). Множество действительных чисел, в отличие от множества рациональных чисел, имеет еще одно, существенно новое свойство—свойство непрерывности.

Существует несколько различных формулировок свойства непрерывности множества действительных чисел. Одна из них будет приведена ниже и называется *принципом вложенных отрезков* или *аксиомой непрерывности* множества действительных чисел (по Кантору).

Если заданы два числа a и b, $(a \leqslant b)$, то множество всех чисел x таких, что $a \leqslant x \leqslant b$, называется *числовым отрезком* и обозначается $[a; b]$. Число $b - a$ называется *длиной* числового отрезка.

Множество числовых отрезков

$$[a_1; b_1], [a_2; b_2], \ldots, [a_n; b_n]$$

называется *системой вложенных отрезков*, если

$$a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n \leqslant b_n \leqslant b_{n-1} \leqslant \ldots \leqslant b_1.$$

Относительно системы вложенных отрезков $[a_n; b_n] (n = 1, 2, 3, \ldots)$ говорят, что длина этих отрезков стремится к нулю с возрастанием n, если для любого числа $\varepsilon > 0$ существует такой номер n_0, что для всех номеров $n \geqslant n_0$ выполняется неравенство

$$b_n - a_n < \varepsilon.$$

Принцип вложенных отрезков. *Для всякой системы вложенных отрезков существует хотя бы одно число, которое принадлежит всем отрезкам данной системы.*

Из принципа вложенных отрезков, в частности, вытекает, что для всякой системы вложенных отрезков, по длине стремящихся к нулю, существует единственное число, принадлежащее всем отрезкам данной системы.

5.3. Представление действительных чисел десятичными дробями. По свойству Архимеда для любого неотрицательного числа a найдется такое неотрицательное целое число N, что

$$N \leqslant a < N+1.$$

Разобьем числовой отрезок $[N;\ N+1]$ на десять равных частей и рассмотрим отрезки

$$[N;\ N,\ 1],\ [N,\ 1;\ N,\ 2],\ \ldots,\ [N,\ 9;\ N+1].$$

Возможны два случая: либо a не совпадает ни с одной из точек деления, либо a является одной из точек деления. В первом случае a принадлежит только одному из перечисленных отрезков, который обозначим I_1:

$$I_1 = [N,\ n_1;\ N,\ n_1+1],$$

где n_1 — одна из цифр 0, 1, 2, ..., 9. Если a — точка деления, то в качестве отрезка I_1 выберем тот, для которого a является левым концом.

Разобьем отрезок I_1 на десять равных частей и выберем из полученных десяти отрезков тот, который содержит точку a и для которого a не является правым концом. Обозначим выбранный отрезок через I_2. Продолжая этот процесс, получим систему вложенных отрезков

$$I_k = [N,\ n_1 n_2 \ldots n_k;\ N,\ n_1 n_2 \ldots (n_k+1)],$$

где n_i $(i=1,\ 2,\ 3,\ \ldots,\ k)$ — одна из цифр 0, 1, 2, ..., 9. Каждый из отрезков I_k содержит точку a, и ни для какого из этих отрезков точка a не является правым концом.

Дроби

$$N, n_1 n_2 \ldots n_k \quad \text{и} \quad N, n_1 n_2 \ldots (n_k+1)$$

называют соответственно *нижней* и *верхней подходящими десятичными дробями* порядка k и обозначают $\underline{a}_k$ и $\overline{a}_k$. Эти дроби удовлетворяют соотношениям

$$\begin{aligned} &\underline{a}_k < \overline{a}_k, \\ &\underline{a}_k \leqslant \underline{a}_{k+1}, \\ &\overline{a}_k \geqslant \overline{a}_{k+1}, \\ &\overline{a}_k - \underline{a}_k = \frac{1}{10^k}. \end{aligned} \tag{1}$$

Последовательности $(\underline{a}_k)$ и $(\overline{a}_k)$, образованные соответственно из нижних и верхних подходящих дробей, имеют пределы $\underline{a}$ и $\overline{a}$:

$$\lim_{k\to\infty} N, n_1 n_2 \ldots n_k = \underline{a}; \qquad \lim_{k\to\infty} N, n_1 n_2 \ldots (n_k+1) = \overline{a},$$

В силу последнего из соотношений (1) десятичные дроби $\underline{a}$ и $\overline{a}$ равны, т. е. представляют собой одну и ту же десятичную дробь. С другой стороны, a — единственная точка, принадлежащая всем отрезкам I_k с длинами, стремящимися к нулю.

Таким образом, числу a поставлена в соответствие десятичная дробь $N, n_1 n_2 \ldots n_k \ldots$, которую называют *десятичной записью* (или *десятичным представлением*) числа a, и пишут

$$a = N, n_1 n_2 \ldots n_k \ldots$$

Соответствие между множеством всех действительных чисел и множеством всех десятичных дробей не является взаимно однозначным: разным десятичным дробям может соответствовать одно число. Именно, дробям вида

$$N, n_1 n_2 \ldots n_k\,(9) \quad \text{и} \quad N, n_1 n_2 \ldots (n_k + 1)$$

соответствует одно и то же рациональное число. Так, рациональное число 1/4 допускает запись в виде двух различных десятичных дробей: 0,25 и 0,24 (9), в чем нетрудно убедиться, используя алгоритм перевода бесконечной периодической дроби в рациональную дробь (см. п. 4.5 настоящей главы).

Если рассматривать множество десятичных дробей, не имеющих периода, состоящего только из цифры 9 (такие дроби называют *допустимыми*), то между множеством всех действительных чисел и множеством всех допустимых десятичных дробей можно установить взаимно однозначное соответствие и отождествить само число с его десятичным представлением.

Исходя из определений суммы, разности, произведения и частного двух рациональных чисел, с помощью представления действительного числа в виде десятичной дроби можно ввести понятия суммы, произведения, разности и частного любых двух действительных чисел (кроме деления на нуль), а также понятие абсолютной величины действительного числа.

Пусть a и b — два действительных числа:

$$a = N_0, n_1 n_2 \ldots n_k \ldots; \quad b = M_0, m_1 m_2 \ldots m_k \ldots$$

Суммой, *произведением* и *разностью* двух действительных чисел a и b называются числа

$$\lim_{k\to\infty} (\underline{a}_k + \underline{b}_k), \quad \lim_{k\to\infty} (\underline{a}_k \cdot \underline{b}_k), \quad \lim_{k\to\infty} (\underline{a}_k - \underline{b}_k),$$

обозначаемые соответственно $a + b$, $a \cdot b$ и $a - b$.

Частным двух чисел a и b при $b \neq 0$ называется число

$$\lim_{k\to\infty} \frac{\underline{a}_k}{\underline{b}_k},$$

обозначаемое $\dfrac{a}{b}$.

При таком определении частного двух действительных чисел может оказаться, что $b_k = 0$ при некоторых k. Однако в силу условия $b \neq 0$ всегда найдется такой номер k_0, что $b_k \neq 0$ при $k \geqslant k_0$, и тогда последовательность $\underline{a}_k/\underline{b}_k$ следует рассматривать не при всех $k \in \mathbf{N}$, а при $k = k_0, k_0 + 1, k_0 + 2, \ldots$

Абсолютной величиной (*модулем*) действительного числа a называется неотрицательное число

$$|a| = \lim_{k \to \infty} |a_k|.$$

5.4. Геометрическое изображение множества действительных чисел. Рассмотрим прямую линию (которую будем называть *осью*) с выбранной на ней произвольной точкой, которую обозначим буквой O (пока именно буквой, а не числом 0). Выберем любую другую точку E, лежащую на прямой справа от точки O. Точка O делит прямую

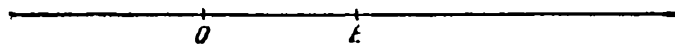


Рис. 2.1

на два луча: луч, которому принадлежит точка E и который называют *положительной* полуосью, и второй луч, который называют *отрицательной полуосью*.

Рассмотрим теперь отрезок прямой с концами O и E (рис. 2.1). Длину отрезка OE, обозначаемую $|OE|$, примем за единицу длины (отрезок OE называют *масштабным* отрезком). Установим теперь соответствие между двумя числами 0 и 1 и точками O и E, полагая, что точка, обозначенная буквой O, соответствует числу 0, а точка, обозначенная буквой E, — числу 1 (точку O называют *началом координат*). Теперь легко установить соответствие между натуральными числами и точками прямой. Например, точка, соответствующая числу 5, будет лежать справа от начала координат на расстоянии, равном пяти длинам отрезка OE; отрицательному целому числу —3 будем ставить в соответствие точку, лежащую слева от точки O на расстоянии, равном трем длинам отрезка OE: $3 \cdot |OE|$.

Далее, если взять единичный отрезок OE, разделить его на n равных частей и отложить от точки O вправо n-ю часть отрезка OE, то правый конец отложенного отрезка считается точкой, изображающей число $1/n$. Если отложить вправо от точки O сумму m отрезков длины $|OE|/n$, то правый конец суммы есть точка, изображающая положительное рациональное число m/n. Аналогично можно построить точки, соответствующие отрицательным рациональным числам.

Остается указать способ сопоставления иррациональных чисел с точками прямой. Пусть положительное число a иррационально. Построим последовательности подходящих десятичных дробей для иррационального числа a:

$$\underline{a}_1;\ \underline{a}_2;\ \underline{a}_3;\ \ldots;\ \underline{a}_k;\ \ldots \quad \text{и} \quad \overline{a}_1;\ \overline{a}_2;\ \overline{a}_3;\ \ldots;\ \overline{a}_k;\ \ldots$$

Члены этих последовательностей — рациональные числа, которым отвечают соответствующие точки прямой. Построим систему отрезков $[\underline{a}_1; \overline{a}_1]$; $[\underline{a}_2; \overline{a}_2]$; $\ldots$; $[\underline{a}_k; \overline{a}_k]$; $\ldots$, вложенных друг в друга. Длина отрезков стремится к нулю с возрастанием k. Эти отрезки имеют только одну общую точку, а именно точку, отвечающую иррациональному числу a.

Таким образом, можно отобразить множество всех действительных чисел $\mathbf{R}$ на множество точек прямой. Верно и обратное: каждой точке прямой можно поставить в соответствие единственное действительное число.

Между множеством действительных чисел и множеством точек прямой существует взаимно однозначное соответствие, и часто вовсе не различают эти два множества, говоря просто «числовая прямая».

5.5. Степени и корни. Свойство ассоциативности операции умножения действительных чисел позволяет ввести понятие натуральной степени любого действительного числа. n-й натуральной *степенью* действительного числа a называется действительное число b, полученное в результате умножения числа a на себя n раз:

$$b = \underbrace{a \cdot a \cdot a \cdot \ldots \cdot a}_{n \text{ сомножителей}};$$

n-ю степень числа a обозначают a^n и пишут

$$b = a^n.$$

Число a называется *основанием степени*, а число n — *показателем степени*.

При $a \neq 0$ по определению $a^0 = 1$; 0^0 не определен.

При $a \neq 0$ по определению $a^{-1} = 1/a$.

Корнем n-й степени (или *арифметическим корнем* n-й степени) из положительного действительного числа a называется единственное положительное решение уравнения $x^n = a$.

Корень n-й степени из числа a обозначается символом $\sqrt[n]{a}$ или $a^{1/n}$ *). Корень второй степени (*квадратный корень*) обычно обозначают просто $\sqrt{a}$.

При $a = 0$ $\sqrt[n]{0} = 0$ (или $0^{1/n} = 0$).

Если действительное число a отрицательно, то корень n-й степени из числа a определяется лишь при нечетном n как единственное действительное отрицательное решение уравнения $x^n = a$.

Корень n-й степени из отрицательного действительного числа a обозначается тем же символом: $\sqrt[n]{a}$ (n — нечетное).

Рациональной степенью m/n (m — целое, n — натуральное) положительного действительного числа a называется число $\left(\sqrt[n]{a}\right)^m$. Рациональная степень числа a также записывается в виде

$$\left(\sqrt[n]{a}\right)^m = (a^{1/n})^m = (a^m)^{1/n} = \sqrt[n]{a^m} = a^{m/n}.$$

Из определения рациональной степени положительного числа следуют равенства

$$\sqrt[m]{a} \cdot \sqrt[n]{a} = \sqrt[m \cdot n]{a^{m+n}}, \quad \sqrt[m]{\sqrt[n]{a}} = \sqrt[m \cdot n]{a},$$

$$\sqrt[n]{a \cdot b} = \sqrt[n]{a} \cdot \sqrt[n]{b}, \quad \sqrt[n]{\frac{a}{b}} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}}.$$

Степень с любым действительным показателем может быть введена, например, следующим образом. Пусть b — действительное число, записанное в виде бесконечной десятичной дроби:

$$b = N, n_1 n_2 \ldots n_k \ldots,$$

*) Символ $\sqrt[n]{\ }$ иногда называют *радикалом* n-й степени.

и пусть $(\underline{b}_k)$ — последовательность нижних подходящих десятичных дробей k-го порядка для числа b, т. е.

$$\underline{b}_1 = N,\ n_1;\quad \underline{b}_2 = N,\ n_1 n_2;\quad \ldots;\quad \underline{b}_k = N,\ n_1 n_2 \ldots n_k;\quad \ldots$$

Для любого положительного числа a можно образовать бесконечную последовательность

$$a^{\underline{b}_1};\quad a^{\underline{b}_2};\quad \ldots;\quad a^{\underline{b}_k};\quad \ldots$$

Предел этой последовательности обозначается a^b и представляет собой действительную степень числа a.

Из определения числа a^b следуют равенства

$$a^{\alpha+\beta} = a^\alpha \cdot a^\beta,\quad (a^\alpha)^\beta = a^{\alpha\cdot\beta},\quad (a\cdot b)^\alpha = a^\alpha \cdot b^\alpha,$$
$$a^\alpha / a^\beta = a^{\alpha-\beta},\quad (a/b)^\alpha = a^\alpha / b^\alpha.$$

5.6. Логарифмы. Пусть a — положительное действительное число, отличное от единицы, а M — любое положительное действительное число. *Логарифмом числа M по основанию a* называется число, обозначаемое $\log_a M$, такое, что

$$a^{\log_a M} = M.$$

Логарифм $\log_a M$ положительного числа M при положительном (отличном от единицы) основании a можно также определить как решение уравнения

$$a^x = M.$$

Основные свойства логарифмов:

$$\log_a a = 1 \quad (a > 0,\ a \neq 1);\quad \log_a (a^k) = k \quad (a > 0,\ a \neq 1);$$
$$\log_a (M_1 \cdot M_2) = \log_a M_1 + \log_a M_2 \quad (M_1 > 0,\ M_2 > 0);$$
$$\log_a (M_1 / M_2) = \log_a M_1 - \log_a M_2 \quad (M_1 > 0,\ M_2 > 0);$$
$$\log_a (b^c) = c \log_a b;\quad \log_a c = \frac{\log_b c}{\log_b a};\quad \log_a c = \frac{1}{\log_c a}.$$

Десятичные логарифмы (логарифмы по основанию 10) $\log_{10} a$ обычно обозначают $\lg a$.

Натуральные логарифмы (т. е. логарифмы по основанию $e = 2{,}718281828459045\ldots$) $\log_e a$ обычно обозначают $\ln a$.

Десятичные и натуральные логарифмы связаны равенствами

$$\ln a = \frac{\lg a}{\lg e} = \ln 10 \cdot \lg a = (2{,}30259\ldots) \lg a,$$
$$\lg a = \frac{\ln a}{\ln 10} = \ln a \cdot \lg e = (0{,}43429\ldots) \ln a.$$

5.7. Пропорции. Равенство вида

$$\frac{a}{b} = \frac{c}{d} \tag{2}$$

называют *пропорцией*, a, b, c, d — *членами* пропорции. В пропорции (2) a и d называют *крайними*, а b и c — *средними* членами пропорции.

Из равенства (2) следует:

1) $a \cdot d = b\ c$;

2) $\frac{a \pm b}{b} = \frac{c \pm d}{d}$, $\frac{a-b}{a+b} = \frac{c-d}{c+d}$; $\frac{ma+nb}{pa+qb} = \frac{mc+nd}{pc+qd}$ (*производные пропорции*), где m, n, p, q — произвольные числа (p и q не равны нулю одновременно).

5.8. Десятичные представления иррациональных чисел. Рациональное число определяется как число, которое может быть записано в виде m/n, где m и n — целые числа и $n \neq 0$ (см. п. 3.2).

Иррациональным числом называется число, которое нельзя представить в виде m/n, где m, n — целые числа и $n \neq 0$.

Любое рациональное число может быть записано либо в виде конечной десятичной дроби, либо в виде бесконечной периодической десятичной дроби (см. п. 4.6). Так как любое действительное число может быть записано в виде десятичной дроби (см. п. 5.3), то любое иррациональное число может быть записано в виде бесконечной непериодической десятичной дроби.

Это свойство иррациональных чисел иногда принимают за определение иррациональных чисел.

Иррациональным числом называется действительное число, десятичная запись которого является бесконечной непериодической десятичной дробью.

Свойства иррациональных чисел. В противоположность множеству рациональных чисел, которое было замкнуто относительно операций сложения, вычитания, умножения и деления (исключая деление на нуль), множество иррациональных чисел не обладает свойством замкнутости ни для одной из перечисленных выше операций. Чтобы убедиться в незамкнутости множества иррациональных чисел, например относительно операции сложения, достаточно указать лишь одну пару иррациональных чисел, сумма которых рациональна. В качестве пары таких чисел могут быть взяты, например, числа 0,1010010001 ...; 0,0101101110 ..., первое из которых образовано последовательностью единиц, разделенных соответственно одним нулем, двумя нулями, тремя нулями и т. д., второе — последовательностью нулей, между которыми поставлены одна единица, две единицы, три единицы и т. д. Каждое из этих чисел будет иррациональным числом, в то время как их сумма будет рациональным числом, десятичное представление которого имеет вид

$$0{,}111 \ldots 1 \ldots = 0{,}(1) = 1/9.$$

Сумма, разность, произведение и частное иррационального числа a и рационального числа r — иррациональные числа.

Из этого свойства иррациональных чисел, в частности, следует, что, имея лишь одно иррациональное число, с помощью рациональных чисел можно получить бесконечно много иррациональных чисел.

5.9. Некоторые способы доказательства иррациональности чисел. Доказательство иррациональности чисел, основанное на определении рационального числа. Иррациональность некоторых чисел может быть доказана с помощью метода доказательства «от противного».

Пусть, например, требуется доказать иррациональность числа $\sqrt{2}$. Предположим, что $\sqrt{2}$ — рациональное число, т. е. может быть

представлено в виде

$$\sqrt{2} = m/n, \tag{3}$$

где m и n — взаимно простые натуральные числа. Для доказательства невозможности представления числа $\sqrt{2}$ в виде (3) мы воспользуемся тем, что числа m и n — взаимно простые. Говоря точнее, мы воспользуемся тем, что числа m и n не являются оба четными — в противном случае дробь m/n была бы сократимой. Возведем обе части равенства (3) в квадрат:

$$2 = \frac{m^2}{n^2} \Leftrightarrow 2 \cdot n^2 = m^2.$$

Число $2 \cdot n^2$ четно. Поэтому m^2 и, следовательно, m также четно. Положив $m = 2 \cdot k$, равенство $2 \cdot n^2 = m^2$ можно записать в виде

$$2 \cdot n^2 = (2 \cdot k)^2 \Leftrightarrow 2 \cdot n^2 = 4 \cdot k^2 \Leftrightarrow n^2 = 2 \cdot k^2.$$

Из последнего равенства видно, что число n^2 также оказывается четным; следовательно, n четно. Мы пришли к заключению, что m и n — четные числа, в то время как дробь m/n по предположению несократима. Полученное противоречие и доказывает, что число $\sqrt{2}$ непредставимо в виде m/n и, следовательно, иррационально.

Аналогичным способом может быть доказана и иррациональность числа $\sqrt{3}$. Только в отличие от предыдущего случая, здесь доказательство будет основано на том, что числа m и n, предполагаемые взаимно простыми, окажутся оба кратными трем.

Доказательство иррациональности чисел с помощью основной теоремы арифметики. Иррациональность некоторых чисел можно доказать с помощью основной теоремы арифметики (см. п. 1.3). По основной теореме арифметики любое натуральное число единственным образом разлагается в произведение простых сомножителей.

Например, докажем, что число $\lg 2$ иррационально. Предположим, что это не так, т. е. найдутся такие натуральные числа m и n, что

$$\lg 2 = m/n, \quad \text{или} \quad 2 = 10^{m/n}.$$

Возведем обе части последнего равенства в степень n:

$$2^n = 10^m \Leftrightarrow 2^n = 2^m \cdot 5^m.$$

Согласно основной теореме арифметики равенство $2^n = 2^m \cdot 5^m$ невозможно, так как 2^n ни при каких n не делится на 5, в то время как $2^m \cdot 5^m$ делится на 5 (поскольку m — натуральное число). Следовательно, число $\lg 2$ иррационально.

Один общий метод доказательства иррациональности чисел (в том числе иррациональности некоторых значений тригонометрических функций). Этот метод основан на следующей теореме:

Если алгебраическое уравнение

$$n_0 x^k + n_1 x^{k-1} + \ldots + n_{k-1} x + n_k = 0$$

с целыми коэффициентами имеет рациональный корень m/n (числа m и n взаимно простые), то число m является делителем числа n_k, а число n — делителем числа n_0.

В частности, рациональными корнями уравнения

$$x^k + n_1 x^{k-1} + \ldots + n_{k-1} x + n_k = 0$$

с целыми коэффициентами и старшим коэффициентом, равным единице, могут быть лишь целые числа.

Доказательство иррациональности чисел с помощью сформулированной теоремы проводится следующим образом *).

1. Составляют алгебраическое уравнение наименьшей натуральной степени с целыми коэффициентами, одним из корней которого заведомо будет число a, иррациональность которого необходимо доказать.

2. Находят все простые делители первого коэффициента n_0 полученного уравнения и свободного члена n_k. Из полученных целых чисел составляют всевозможные рациональные числа с числителем, который является делителем числа n_k, и знаменателем, который является делителем числа n_0. Тогда по сформулированной выше теореме лишь эти рациональные числа могут быть рациональными корнями данного уравнения.

3. Подстановкой этих рациональных чисел в уравнение проверяют, является ли хотя бы одно из них корнем уравнения. Если ни одно из построенных рациональных чисел не является корнем уравнения, то данное уравнение рациональных корней не имеет и, следовательно, рассматриваемое число иррациональное. Если же какое-нибудь из построенных рациональных чисел является корнем уравнения, то доказывают, что рассматриваемое число не равно этому корню и, следовательно, оно иррационально.

Примеры. 1. Доказать, что число $\sqrt[3]{2}-\sqrt{3}$ иррационально.

Положим $x=\sqrt[3]{2}-\sqrt{3}$. Тогда $x+\sqrt{3}=\sqrt[3]{2}$. Возведем обе части равенства в куб; после несложных преобразований получим уравнение

$$x^3+9x-2=-3\sqrt{3}(x^2+1).$$

Возведя обе части полученного уравнения в квадрат, после приведения подобных членов придем к уравнению

$$x^6-9x^4-4x^3+27x^2-36x-23=0.$$

Из способа построения этого уравнения следует, что число $\sqrt[3]{2}-\sqrt{3}$ — его корень. С другой стороны, единственно возможными рациональными корнями этого уравнения являются целые числа — делители числа -23, т. е. $+1$, -1, $+23$, -23. Непосредственная подстановка этих чисел в уравнение показывает, что эти числа не являются корнями уравнения. Таким образом, рациональных корней уравнение не имеет и значит, число $\sqrt[3]{2}-\sqrt{3}$ иррационально.

2. Доказать, что число $\cos 20^\circ$ иррационально.

По формуле тройного аргумента $\cos 60^\circ=1/2$ и $\cos 20^\circ$ связаны формулой

$$\cos 60^\circ=4\cos^3 20^\circ-3\cos 20^\circ.$$

Обозначив $\cos 20^\circ=x$, запишем последнее равенство в виде

$$8x^3-6x-1=0. \qquad (4)$$

Число $x=\cos 20^\circ$ — корень этого уравнения. Применяя к этому уравнению сформулированную выше теорему, видим, что возможными рациональными корнями этого уравнения являются числа ± 1, $\pm 1/2$, $\pm 1/4$, $\pm 1/8$. Подстановкой этих чисел в уравнение (4) можно убедиться, что ни одно из них не является его корнем. Следовательно, уравнение (4) не имеет рациональных корней, поэтому число $\cos 20^\circ$ иррационально. Следует отметить, что подобный способ непосредственного использования формулы тройного аргумента непригоден для доказательства, например, иррациональности числа $\cos 10^\circ$, так как алгебраическое уравнение третьей степени, корнем которого будет число $x=\cos 10^\circ$, имеет иррациональные коэффициенты.

*) Излагаемый ниже метод применим только для доказательства иррациональности алгебраических чисел (см. п. 5.10).

Иррациональность числа $\cos 10^\circ$ проще доказать в два этапа: сначала доказать, что число $\cos 20^\circ$ иррационально, после чего, воспользовавшись формулой

$$1+\cos 20^\circ = 2\cos^2 10^\circ,$$

методом от противного доказать, что число $\cos 10^\circ$ не может быть рациональным.

Сформулируем еще одно утверждение, позволяющее доказать иррациональность некоторых значений тригонометрических функций.

Если угол θ таков, что число $\cos 2\theta$ иррационально, то числа $\cos\theta$, $\sin\theta$, $\operatorname{tg}\theta$ также иррациональны. Чтобы убедиться в этом, достаточно применить метод доказательства от противного, предварительно воспользовавшись тригонометрическими формулами половинного аргумента и известным тригонометрическим тождеством

$$1+\operatorname{tg}^2\theta = \frac{1}{\cos^2\theta}.$$

Так, предположим, что $\sin\theta$ — число рациональное; тогда $2\sin^2\theta$ будет тоже рациональным числом. Так как $\sin^2\theta$ и $\cos 2\theta$ связаны равенством

$$1-\cos 2\theta = 2\sin^2\theta,$$

из рациональности числа $2\sin^2\theta$ будет следовать рациональность числа $\cos 2\theta$. Полученное противоречие и доказывает, что число $\sin\theta$, как и число $\cos 2\theta$, иррационально.

5.10. Алгебраические и трансцендентные числа. Множество действительных чисел, помимо разбиения на множество рациональных и множество иррациональных чисел, может быть разбито на два других множества — множество алгебраических чисел и множество трансцендентных чисел.

Действительное число, являющееся корнем некоторого алгебраического уравнения вида

$$a_0x^k + a_1x^{k-1} + \ldots + a_{k-1}x + a_k = 0$$

с целыми коэффициентами $a_0, a_1, \ldots, a_{k-1}, a_k$, из которых хотя бы один отличен от нуля, называется *алгебраическим числом*.

Действительное число, не являющееся корнем никакого алгебраического уравнения с целыми коэффициентами, называется *трансцендентным числом*.

Множество всех рациональных чисел является подмножеством множества алгебраических чисел, так как любое рациональное число m/n является корнем алгебраического уравнения первой степени

$$nx - m = 0$$

Отсюда также следует, что всякое трансцендентное число иррационально.

Разбиение множества действительных чисел на различные подмножества схематически можно изобразить так:

Действительные числа — Рациональные (все они являются алгебраическими); Иррациональные — Алгебраические, Трансцендентные

Действительные числа — Алгебраические — Рациональные, Иррациональные; Трансцендентные (все они являются иррациональными)

Множество всех алгебраических чисел образует счетное множество, а множество всех трансцендентных чисел — несчетное.

Некоторые трансцендентные числа. Ранее (см. п. 5.9) были приведены доказательства иррациональности некоторых чисел, в частности иррациональности числа $\lg 2$. В то же время $\lg 2$ является трансцендентным

числом. Впервые предположение о трансцендентности числа lg 2 было высказано Л. Эйлером в XVIII веке. Доказательство трансцендентности числа lg 2 значительно сложнее, чем доказательство его иррациональности, и основано на методах гораздо более общих и глубоких, чем те, которые используются при доказательстве его иррациональности. Трансцендентность числа lg 2 следует из теоремы.

Число a^b трансцендентно, если числа a и b алгебраические (случаи $a = 0$, $a = 1$ и случай рационального b исключены).

Трансцендентными числами будут и хорошо известные числа e и π; их трансцендентность была доказана в конце XIX века. Вопрос об алгебраичности или трансцендентности числа π связан с решением одной геометрической задачи, сформулированной еще за несколько веков до нашей эры. Эта задача называется задачей о квадратуре круга и формулируется так: можно ли с помощью только циркуля и линейки построить квадрат с площадью, равной площади круга единичного радиуса?

Доказано, что с помощью циркуля и линейки можно строить лишь такие геометрические фигуры, площадь которых выражается алгебраическим числом. Поэтому трансцендентность числа π указывает на невозможность решения задачи о квадратуре круга.

Напротив, значения тригонометрических функций при некоторых значениях аргумента (например, значения $\cos 20°$ и $\sin 10°$, иррациональность которых была доказана в п. 5.9) являются алгебраическими числами, что вытекает из следующей теоремы.

Для любого рационального числа r числа $\sin(90°\cdot r)$, $\cos(90°\cdot r)$ и $\lg(90°\cdot r)$ являются алгебраическими.

§ 6. Приближенные вычисления

6.1. Приближенное значение числа и погрешности. Для выполнения вычислений действительные числа часто бывает удобно записывать в виде десятичных дробей. Однако не любое рациональное, а тем более действительное число может быть записано в виде конечной дроби, с которой можно проводить требуемые вычисления. Поэтому часто бесконечную десятичную дробь, представляющую запись данного действительного числа, приходится обрывать на некоторой цифре после запятой и оперировать с полученной конечной десятичной дробью. Число, представляемое конечной десятичной дробью, полученной в результате обрывания бесконечной десятичной дроби, называется *приближенным значением* данного числа. Например, различными приближенными значениями числа $\pi = 3{,}14159\ldots$ будут числа 3,14; 3,1415 и т. д.; в качестве приближенных значений числа $e = 2{,}71828\ldots$ можно взять числа 2,7; 2,71; 2,7182 и т. д. Приближенные значения чисел могут появиться не только в результате обрыва записи числа в виде десятичной дроби. При решении любой практической задачи мы используем числа, полученные в результате некоторых измерений. Однако при всяких измерениях появляется ошибка, связанная с несовершенством наших измерительных инструментов, геометрической неправильностью измеряемого предмета и т. п.

Обозначим приближенное значение числа a через a^*. Абсолютную величину разности числа a и его приближенного значения a^*

$$|a - a^*|$$

называют *абсолютной погрешностью* приближенного значения a^*.

Абсолютная погрешность является одной из характеристик приближенного значения числа. Часто она представляет только теоре-

тический интерес, так как в большинстве случаев числа a мы можем и не знать. В ряде случаев можно указать границы, в которых изменяется абсолютная погрешность. Эти границы обычно определяются способом получения приближенного значения числа. Например, производя измерения длины с помощью линейки с метками, расставленными на расстоянии в 1 мм, мы можем гарантировать, что абсолютная погрешность измерений не будет превышать 1 мм.

Предельной абсолютной погрешностью приближенного значения числа называют такое число A, что

$$|a-a^*| < A.$$

По своему определению предельная абсолютная погрешность может вводиться неоднозначно. Так, в случае рассмотренного выше примера измерения длин линейкой предельная абсолютная погрешность может быть взята равной 1 мм, 1,5 мм и т. д.

Заметим, что ни абсолютная погрешность, ни предельная абсолютная погрешность сами по себе еще не характеризуют, с какой точностью мы заменяем число a его приближенным значением. Например, пусть предельная погрешность результата некоторого измерения равна 10 км. Если измерялось расстояние между двумя домами одного и того же города, то такая точность, вообще говоря, неудовлетворительна. Если же измерялось расстояние между полюсами Земли, то указанная точность очень хороша.

Характеристиками точности результата измерения (или замены числа a его приближенным значением), в которых кроме погрешности участвует и сам результат измерений, являются величины, называемые относительной погрешностью и предельной относительной погрешностью.

Относительной погрешностью называется отношение абсолютной погрешности к абсолютной величине приближенного значения числа:

$$\delta = \frac{|a-a^*|}{|a^*|}.$$

Предельной относительной погрешностью приближенного значения числа называется отношение предельной абсолютной погрешности к абсолютной величине приближенного значения числа:

$$\Delta = \frac{A}{|a^*|}.$$

Например, число 3,14 является приближенным значением числа π. Абсолютная погрешность этого приближенного значения равна 0,00159...; предельную абсолютную погрешность можно считать равной 0,0016, а предельную относительную погрешность

$$\frac{0,0016}{3,14} = 0,000509...,$$

можно положить равной 0,00051.

6.2. Десятичная запись приближенных значений числа. Любое положительное действительное число a может быть записано в виде конечной или бесконечной десятичной дроби:

$$a = n_l \cdot 10^l + n_{l-1} \cdot 10^{l-1} + \ldots + n_{l-k+1} \cdot 10^{l-k+1} + \ldots,$$

где $n_l, n_{l-1}, \ldots, n_{l-k+1}, \ldots$ — цифры числа a, причем $n_l \neq 0$, а l — некоторое целое число.

Приближенное значение a^* числа a можно получить, обрывая десятичную запись числа a. При этом в записи приближенного значения a^* в десятичной позиционной системе счисления иногда приходится вводить дополнительные нули в начале или в конце числа. Например, в качестве приближенного значения числа

$$b = 2\cdot 10^{-2} + 4\cdot 10^{-3} + 0\cdot 10^{-4} + 9\cdot 10^{-5} + \ldots$$

можно взять число

$$b^* = 0,0\,2409,$$

а в качестве приближенного значения числа

$$c = 1\cdot 10^5 + 0\cdot 10^4 + 7\cdot 10^3 + 3\cdot 10^2 + \ldots$$

можно взять число

$$c^* = 107\,000.$$

Значащей цифрой приближенного значения числа называется всякая отличная от нуля цифра его десятичной записи, а также нуль, если он находится между значащими цифрами или является представителем сохраненного десятичного разряда.

Нули, стоящие перед и после запятой в записи числа b^*, и нули, стоящие в конце записи числа c^*, не считаются значащими цифрами.

Пусть, например, число 0,002080 есть приближенное значение некоторого числа. Первые три нуля не являются значащими цифрами, так как они служат только для указания десятичных разрядов других цифр. Остальные два нуля являются значащими цифрами, так как первый из них находится между значащими цифрами 2 и 8, а второй указывает, что приближенное значение числа получено обрыванием десятичной записи числа с сохранением шестого разряда после запятой. В случае, если в приближенном значении 0,002080 последняя цифра не является значащей, оно должно быть записано в виде 0,00208. Таким образом, приближенные значения 0,002080 и 0,00208 не равноценны, так как первое из них содержит четыре значащих цифры, а второе — три.

При записи приближенных значений целых чисел нули, стоящие в конце десятичной записи числа, могут служить как обозначения значащих цифр, так и для указания разряда других цифр. Поэтому обычная запись приближенных значений целых чисел может пониматься неоднозначно. Например, пусть известно, что число 134 000 есть приближенное значение некоторого числа. По виду числа 134 000 нельзя сделать вывод, сколько значащих цифр в этом числе (хотя сразу можно сказать, что их не меньше трех). Для того чтобы избежать такой неопределенности, целое число записывают с помощью десятичных дробей и степеней числа 10. Например, если хотят указать, что число 134 000 имеет три значащих цифры, то его записывают в виде $0,134\cdot 10^6$ (или $1,34\cdot 10^5$, или $13,4\cdot 10^4$, или $134\cdot 10^3$). Если же хотят указать, что число 134 000 имеет пять значащих цифр, то его записывают в виде $0,13400\cdot 10^6$ (или соответственно $1,3400\cdot 10^5$, $13,400\cdot 10^4$, $134,00\cdot 10^3$).

Говорят, что n первых значащих цифр приближенного значения числа являются *верными*, если абсолютная погрешность приближенного значения числа не превышает половины единицы разряда n-й значащей цифры, считая разряды слева направо.

Таким образом, если для приближенного значения a^* числа

$$a = n_l \cdot 10^l + n_{l-1} \cdot 10^{l-1} + \ldots + n_{l-k+1} \cdot 10^{l-k+1} + \ldots$$

известно, что

$$|a - a^*| \leqslant 0{,}5 \cdot 10^{l-k+1},$$

то первые k цифр n_l, n_{l-1}, ..., n_{l-k+1} приближенного значения являются верными.

Например, для числа 121,57 число 122,00 является приближенным значением с тремя верными цифрами, так как

$$|121{,}57 - 122| = 0{,}43 < 0{,}5 \cdot 10^0.$$

Результат арифметических действий над приближенными значениями чисел представляет собой также приближенное значение некоторого числа. Погрешность полученного результата оценивается с помощью следующих правил.

1. Предельная абсолютная погрешность алгебраической суммы приближенных значений чисел не превышает сумму предельных абсолютных погрешностей слагаемых.

2. Относительная погрешность суммы приближенных значений чисел заключена между наибольшей и наименьшей из относительных погрешностей слагаемых.

В согласии со сформулированными правилами, во избежание лишних вычислений, при подсчете суммы приближенных значений чисел производят предварительное округление слагаемых (см. п. 6.3). При этом слагаемое, имеющее наименьшее число верных знаков, оставляют неокругленным, а в остальных слагаемых оставляют на один или два знака больше по сравнению с неокругляемым числом.

6.3. Округление чисел. Чтобы округлить число до n значащих цифр, отбрасывают все его цифры, стоящие после n-го разряда, или, если это нужно для сохранения разрядов, заменяют их нулями. При этом:

1) если за последней сохраняемой цифрой следует цифра 0, 1, 2, 3 или 4, то при округлении сохраняются все цифры до последней сохраняемой включительно. Такое округление называется *округлением с недостатком*;

2) если за последней сохраняемой цифрой следует 9, 8, 7, 6 или 5, то к последней сохраняемой цифре прибавляется единица. Такое округление называют *округлением с избытком*.

Примеры. 1. Округление числа 219,364 до четырех значащих цифр дает число 219,4. Последняя цифра округленного числа увеличена на единицу, так как первая отбрасываемая цифра больше пяти.

2. Округление числа 6,4502 до двух значащих цифр дает число 6,5. Последняя цифра округленного числа увеличена на единицу, так как первая отбрасываемая цифра равна пяти.

3. Округление числа 1836 до трех значащих цифр дает число $1{,}84 \cdot 10^3$.

При использовании правил округления абсолютная погрешность округления не превосходит половины единицы разряда последней оставленной значащей цифры, т. е. все значащие цифры приближенного значения числа, полученного с помощью правил округления, являются верными цифрами.

6.4. Метод касательных (метод Ньютона). Метод касательных является одним из самых эффективных численных методов решения уравнения.

Допустим, нам известно, что уравнение $f(x)=0$ имеет корень на промежутке $[a; b]$ (функция $f(x)$ дифференцируема на $[a; b]$, и ее производная не обращается на $[a; b]$ в нуль). Отыскать этот корень можно следующим способом. Возьмем произвольную точку $x_1 \in [a; b]$ и напишем уравнение касательной к графику функции $f(x)$ в точке с абсциссой x_1 *):

$$y=f(x_1)+f'(x_1)(x-x_1).$$

Решая уравнение

$$f(x_1)+f'(x_1)(x-x_1)=0,$$

найдем точку пересечения касательной с осью Ox:

$$x=x_1-\frac{f(x_1)}{f'(x_1)}.$$

Обозначим найденную точку через x_2:

$$x_2=x_1-\frac{f(x_1)}{f'(x_1)}.$$

Напишем уравнение касательной к графику функции $f(x)$ в точке с абсциссой x_2:

$$y=f(x_2)+f'(x_2)(x-x_2),$$

и найдем точку x_3 пересечения касательной с осью Ox: $x_3=x_2-\frac{f(x_2)}{f'(x_2)}$.

Продолжая этот процесс, получим последовательность (x_n), задаваемую рекуррентной формулой

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}, \qquad n\in\mathbb{N} \tag{1}$$

Если функция $f(x)$ удовлетворяет некоторым дополнительным условиям (например, $f(x)$ имеет положительную (или отрицательную) вторую производную во всех внутренних точках $[a; b]$), то последовательность (1) сходится к числу, являющемуся корнем уравнения $f(x)=0$.

Метод касательных может быть применен для решения задачи извлечения корня из произвольного положительного числа a. Значение $\sqrt{a}$ будем искать как решение уравнения

$$f(x)=x^2-a=0.$$

Рекуррентная формула (1) в данном случае принимает вид

$$x_{n+1}=\frac{1}{2}\left(x_n+\frac{a}{x_n}\right). \tag{2}$$

Здесь в качестве первого члена последовательности (x_n) можно взять любое число $x_1 \in (0; +\infty)$. Если в качестве числа x_1 выбрано число, удовлетворяющее неравенству $x_1^2-a>0$, то абсолютные погрешности приближенных значений числа $\sqrt{a}$ на n-м и $(n+1)$-м шагах вычислений определяются равенствами

$$\delta_n=x_n-\sqrt{a}>0, \qquad \delta_{n+1}=x_{n+1}-\sqrt{a}>0.$$

Используя рекуррентную формулу (2), установим связь между δ_{n+1} и δ_n:

$$\delta_{n+1}=x_{n+1}-\sqrt{a}=\frac{1}{2}\left(x_n+\frac{a}{x_n}\right)-\sqrt{a}=$$

$$=\frac{1}{2}\left(\delta_n+\sqrt{a}+\frac{a}{\delta_n+\sqrt{a}}\right)-\sqrt{a}=\frac{1}{2}\delta_n+\frac{a}{2(\delta_n+\sqrt{a})}-\frac{1}{2}\sqrt{a}=$$

$$=\frac{1}{2}\left[\delta_n-\sqrt{a}+\frac{a}{\delta_n+\sqrt{a}}\right]=\frac{1}{2}\left[\frac{\delta_n^2}{\delta_n+\sqrt{a}}\right]$$

*) Необходимо, чтобы точка x_1 была расположена достаточно близко от корня уравнения $f(x)=0$, поэтому для выбора точки x_1 иногда приходится проводить предварительное исследование функции $f(x)$.

Поскольку $\delta_n > 0$, то, отбрасывая в знаменателе последней дроби слагаемое δ_n, мы тем самым увеличиваем ее значение, и, следовательно,

$$\delta_{n+1} < \frac{\delta_n^2}{2\sqrt{a}}. \tag{3}$$

Полученная рекуррентная оценка абсолютной погрешности указывает на высокую скорость сходимости последовательности приближенных значений x_n

Например, будем вычислять с помощью метода касательных $\sqrt{33\,520}$. В качестве x_1 возьмем число 200. Это значение данного корня с избытком. Абсолютная погрешность δ_1 не превышает 20. Тогда по формуле (3)

$$\delta_2 < \frac{400}{2\cdot 180} < 1{,}5; \quad \delta_3 < \frac{225}{2\cdot 180} < 0{,}001.$$

Здесь для упрощения оценок $\sqrt{a}$ в формуле (3) заменен на меньшее число 180. Таким образом, при вычислении значения $\sqrt{33\,520}$ уже после третьего шага получается приближенное значение корня, отличающееся от его точного значения меньше чем на 0,001.

6.5. **Алгоритм извлечения квадратного корня из натурального числа.** Рассмотрим на конкретном примере алгоритм извлечения квадратного корня для случая, когда число, стоящее под знаком корня, является квадратом натурального числа. Пусть требуется вычислить $\sqrt{33\,489}$. Число 33 489 разобьем на группы цифр (по две цифры), двигаясь справа налево:

$$3\,34\,89.$$

Ищем наибольшее число, квадрат которого не превосходит числа 3, стоящего в первой группе цифр. Этим числом будет число 1. Записываем его в ответ. Возводим единицу в квадрат и вычитаем из числа 3. К полученной разности приписываем вторую группу цифр. Получаем число 234.

Удваиваем число, которое было записано в ответ (в нашем случае 1), и приписываем к полученному числу справа такую наибольшую цифру, чтобы произведение полученного двузначного числа на эту цифру не превосходило 234. В нашем случае этой цифрой будет цифра 8:

$$28\cdot 8 = 224 < 234.$$

Пишем цифру 8 вслед за цифрой 1 в ответ. Из числа 234 вычитаем число 224 и к полученной разности приписываем последнюю группу цифр. Получаем число 1089.

Удваиваем число, которое было записано в ответ (в нашем случае 18), и приписываем справа к полученному числу (в нашем случае к числу 36) такую наибольшую цифру, чтобы произведение полученного трехзначного числа на приписанную цифру не превосходило числа 1089. В нашем случае этой цифрой будет цифра 3:

$$363\cdot 3 = 1089.$$

Записываем цифру 3 в ответ. Процедура извлечения квадратного корня завершена:

$$\sqrt{33\,489} = 183.$$

Обычно алгоритм извлечения квадратного корня записывают в виде следующей схемы:

$$\begin{array}{r|l} \sqrt{3\,34\,89} = 183. & \\ -1 & \\ \hline -234 & \\ -224 & \\ \hline -1089 & \\ -1089 & \\ \hline 0 & \end{array} \qquad \begin{array}{r} \times 28 \\ 8 \\ \hline 224 \end{array} \qquad \begin{array}{r} \times 363 \\ 3 \\ \hline 1089 \end{array}$$

Глава 3

КОМПЛЕКСНЫЕ ЧИСЛА

Комплексные числа, в отличие от натуральных, целых и рациональных чисел, не являются объектами, используемыми при подсчетах и измерениях. Комплексные числа как некоторые абстрактные математические объекты впервые были использованы при вычислении корней кубических уравнений.

В первой трети XVI века итальянский математик Н. Тарталья показал, что корень кубического уравнения

$$x^3 = px + q$$

может быть вычислен по формуле

$$x = \sqrt[3]{u} + \sqrt[3]{v}, \tag{1}$$

где u и v — решения системы уравнений

$$u + v = q, \quad uv = \left(\frac{p}{3}\right)^3. \tag{2}$$

Например, чтобы найти корень кубического уравнения

$$x^3 = 9x + 28,$$

необходимо составить и решить систему (2) для данного уравнения: $u = 27$, $v = 1$ и $u = 1$, $v = 27$. Используя формулу (1), находим $x = 4$, т. е. число 4 является корнем данного уравнения.

Однако оказалось, что существуют кубические уравнения, для которых система (2) не имеет решений в множестве действительных чисел, в то время как кубическое уравнение заведомо имеет действительный корень. Например, уравнение

$$x^3 = 15x + 4$$

имеет действительный корень $x = 4$, в чем легко убедиться, подставив в данное уравнение вместо x число 4. Если же для данного уравнения написать систему (2), то окажется, что она не имеет решений в множестве действительных чисел. Это явление впервые объяснил итальянский математик Р. Бомбелли в 1572 г., и его объяснение, по существу, было основано на введении понятия комплексного числа и простейших правил действий над комплексными числами. Однако вплоть до XIX века, несмотря на то что аппарат комплексных чисел позволил получить много важных фактов, относящихся также и к действительным числам, математическая сущность комплексных чисел многим крупным математикам представлялась весьма неясной и даже мистической. Лишь в XIX веке после появления работ К. Гаусса и К. Бесселя, в которых давалось наглядное геометрическое изображение комплексных чисел (как точек или векторов на плоскости), а также правила действий над ними, использование комплексных чисел стало общепринятым. Упомянем еще

один факт, который также приводит к мысли о необходимости расширения множества действительных чисел до множества комплексных чисел. Как известно, натуральная степень любого действительного числа есть действительное число. Однако операция извлечения корня (обратная операции возведения в степень) не всегда выполнима в множестве действительных чисел: не существует действительного числа a, четная степень которого была бы отрицательным числом. Другими словами, в множестве действительных чисел не существует числа, которое было бы корнем уравнения

$$x^n = b,$$

где n — четное число, а b — отрицательное действительное число.

Следуя общему плану расширения числовых множеств, как это уже неоднократно делалось (например, при введении понятий отрицательных и рациональных чисел), множество действительных чисел можно расширить до множества чисел, которое будет замкнуто относительно операции извлечения корня.

Один из способов построения множества комплексных чисел заключается в том, что множество действительных чисел расширяется путем присоединения к множеству действительных чисел нового числового объекта — корня уравнения

$$x^2 + 1 = 0.$$

Полученное «расширенное» множество называется *множеством комплексных чисел.*

§ 1. Множество комплексных чисел

1.1. Аксиоматическое построение множества комплексных чисел. Множество комплексных чисел определяется как множество, элементы которого подчиняются перечисленным ниже правилу сравнения I и правилам арифметических действий II, III. Множество всех комплексных чисел обозначается символом C.

Комплексное число обозначается $a + bi$, где a и b — действительные числа, называемые соответственно *действительной частью* и *мнимой частью* комплексного числа $a + bi$, а символ i, определяемый условием $i^2 = -1$, называется *мнимой единицей.*

Обычно комплексное число $a + bi$ обозначают одной буквой (чаще всего z):

$$z = a + bi.$$

Действительную и мнимую части комплексного числа $z = a + bi$ обозначают $\operatorname{Re} z$ и $\operatorname{Im} z$ соответственно *):

$$a = \operatorname{Re} z, \quad b = \operatorname{Im} z.$$

Запись комплексного числа z в виде $a + bi$ называют *алгебраической формой записи комплексного числа.*

I. Два комплексных числа $z_1 = a_1 + b_1 i$ и $z_2 = a_2 + b_2 i$ считаются *равными* ($z_1 = z_2$), если равны их действительные и мнимые части ($a_1 = a_2$; $b_1 = b_2$). Комплексное число $z = a + bi$ считается равным нулю ($z = 0$), если его действительная и мнимая части равны нулю ($a = b = 0$). Комплексное число $z = a + bi$ при $b = 0$ считается совпа-

*) Re — от слова réel — действительный (фр.), Im — от слова Imaginaire — мнимый (фр.).

дающим с действительным числом a ($a+0i=a$). Комплексное число $z=a+bi$ при $a=0$ называется *чисто мнимым* и обозначается bi ($0+bi=bi$).

II. *Суммой* комплексных чисел $z_1=a_1+b_1i$ и $z_2=a_2+b_2i$ называется комплексное число z, действительная часть которого равна сумме действительных частей чисел z_1 и z_2, а мнимая часть — сумме мнимых частей чисел z_1 и z_2, т. е.

$$z=(a_1+a_2)+(b_1+b_2)\,i.$$

О числе z говорят, что оно получено в результате сложения комплексных чисел z_1 и z_2, и пишут

$$z=z_1+z_2.$$

Числа z_1 и z_2 называются *слагаемыми*.

Свойства операции сложения комплексных чисел:

1) ассоциативность: $(z_1+z_2)+z_3=z_1+(z_2+z_3)$;

2) коммутативность: $z_1+z_2=z_2+z_1$.

Комплексное число $-a-bi$ называется *противоположным* комплексному числу $a+bi$. Комплексное число, противоположное комплексному числу z, обозначается $-z$. Сумма комплексных чисел z и $-z$ равна нулю ($z+(-z)=0$).

Разность комплексных чисел $z_1=a_1+b_1i$ и $z_2=a_2+b_2i$ есть комплексное число z, являющееся суммой числа z_1 и числа, противоположного z_2:

$$z=z_1+(-z_2)=(a_1-a_2)+(b_1-b_2)\,i,$$

т. е. комплексное число, действительная и мнимая части которого равны соответственно разности действительных и мнимых частей уменьшаемого и вычитаемого. О числе z говорят, что оно получено в результате вычитания комплексного числа z_2 из комплексного числа z_1, и пишут

$$z=z_1-z_2.$$

III. *Произведением* комплексных чисел $z_1=a_1+b_1i$ и $z_2=a_2+b_2i$ называется комплексное число

$$z=(a_1a_2-b_1b_2)+(a_1b_2+a_2b_1)\,i.$$

О числе z говорят, что оно получено в результате умножения комплексного числа z_1 на комплексное число z_2, и пишут

$$z=z_1\cdot z_2.$$

Числа z_1 и z_2 называются *сомножителями*.

Свойства операции умножения комплексных чисел:

1) ассоциативность: $(z_1\cdot z_2)\cdot z_3=z_1\cdot(z_2\cdot z_3)$;

2) коммутативность: $z_1\cdot z_2=z_2\cdot z_1$.

Частным двух комплексных чисел z_1 и z_2 ($z_2\neq 0$) называется такое комплексное число z, что $z_1=z\cdot z_2$. Частное комплексных чисел $z_1=a_1+b_1i$ и $z_2=a_2+b_2i$ вычисляется по формуле

$$z=\frac{a_1a_2+b_1b_2}{a_2^2+b_2^2}+\frac{a_2b_1-a_1b_2}{a_2^2+b_2^2}\,i.$$

О числе z говорят, что оно получено в результате деления комплексного числа z_1 на комплексное число z_2, и пишут

$$z=\frac{z_1}{z_2} \quad \text{или} \quad z=z_1/z_2.$$

Сложение и умножение комплексных чисел связаны законом дистрибутивности умножения относительно сложения:

$$(z_1+z_2)\,z_3=z_1\cdot z_3+z_2\cdot z_3.$$

Число $\sqrt{a^2+b^2}$ называется *модулем* комплексного числа $z=a+bi$. Модуль комплексного числа обозначают $|z|$. Модули двух любых комплексных чисел z_1 и z_2 удовлетворяют соотношениям (в случае частного предполагается, что $z_2\neq 0$)

$$|z_1+z_2|\leqslant|z_1|+|z_2|, \quad |z_1-z_2|\geqslant||z_1|-|z_2||,$$

$$|z_1\cdot z_2|=|z_1|\cdot|z_2|, \quad |z_1/z_2|=|z_1|/|z_2|, \quad |z^n|=|z|^n=|\bar{z}|^n.$$

Комплексное число $a-bi$ называется *комплексно сопряженным* с числом $z=a+bi$ и обозначается $\bar{z}$.

Свойства комплексно сопряженных чисел

1. Произведение комплексного числа $z=a+bi$ и комплексно сопряженного с ним числа $\bar{z}=a-bi$ есть действительное число:

$$z\cdot\bar{z}=a^2+b^2.$$

2. Если $\bar{z}$ — число, комплексно сопряженное с числом z, то число, комплексно сопряженное с числом $\bar{z}$, есть число z:

$$\overline{(\bar{z})}=z.$$

3. Если z_1, $\bar{z}_1$ и z_2, $\bar{z}_2$ — две пары комплексно сопряженных чисел, то сумма, произведение, разность и частное чисел $\bar{z}_1$ и $\bar{z}_2$ являются числами, комплексно сопряженными сумме, произведению, разности и частному чисел z_1 и z_2 соответственно:

$$\bar{z}_1+\bar{z}_2=\overline{(z_1+z_2)}, \quad \bar{z}_1-\bar{z}_2=\overline{(z_1-z_2)},$$

$$\bar{z}_1\cdot\bar{z}_2=\overline{(z_1\cdot z_2)}, \quad \bar{z}_1/\bar{z}_2=\overline{(z_1/z_2)}, \quad z_2\neq 0.$$

4. Если $z=a+bi$ и $\bar{z}=a-bi$ — пара комплексно сопряженных чисел, то

$$a=\operatorname{Re}z=\frac{z+\bar{z}}{2}, \quad b=\operatorname{Im}z=\frac{z-\bar{z}}{2i}.$$

1.2. Множество комплексных чисел как множество упорядоченных пар действительных чисел. Наряду с аксиоматическим способом существует другой способ построения комплексных чисел.

Рассмотрим множество всевозможных упорядоченных пар действительных чисел $(a;\ b)$, для которых определены сформулированные ниже правило равенства и арифметические операции.

Две пары $(a;\ b)$ и $(c;\ d)$ считают *равными* и пишут

$$(a;\ b)=(c;\ d),$$

если $a=c$ и $b=d$.

Арифметические операции.

Сложение. *Суммой пар* $(a; b)$ и $(c; d)$ называется пара $(a+c; b+d)$. Говорят, что пара $(a+c; b+d)$ получена в результате сложения пар $(a; b)$ и $(c; d)$, и пишут

$$(a; b)+(c; d)=(a+c; b+d).$$

Ассоциативность и коммутативность операции сложения пар следует из определения суммы пар и свойств ассоциативности и коммутативности сложения действительных чисел.

Нулевой парой называют пару $(x; y)$, удовлетворяющую условию

$$(a; b)+(x; y)=(a; b),$$

где $(a; b)$ — произвольная пара. Из определения операции сложения и условия равенства пар следует, что нулевой является пара $(0; 0)$.

Умножение. *Произведением пар* $(a; b)$ и $(c; d)$ называется пара $(ac-bd; ad+bc)$. Говорят, что пара $(ac-bd; ad+bc)$ получена в результате умножения пар $(a; b)$ и $(c; d)$, и пишут

$$(a; b)\cdot(c; d)=(ac-bd; ad+bc).$$

Ассоциативность и коммутативность умножения пар, а также дистрибутивность умножения относительно сложения следуют из определения произведения и суммы пар и соответствующих свойств операций умножения и сложения действительных чисел.

Единичной парой называют пару $(x; y)$, удовлетворяющую условию

$$(a; b)\cdot(x; y)=(a; b),$$

где $(a; b)$ — произвольная пара. Из определения операции умножения пар и условия равенства пар следует, что единичной является пара $(1; 0)$.

Вычитание. *Разностью пар* $(a; b)$ и $(c; d)$ называется пара $(x; y)$, удовлетворяющая равенству

$$(a; b)=(c; d)+(x; y).$$

Из определения операции сложения пар и условия равенства пар следует, что разностью пар $(a; b)$ и $(c; d)$ является пара

$$(x; y)=(a-c; b-d).$$

О паре $(a-c; b-d)$ говорят, что она получена в результате вычитания пары $(c; d)$ из пары $(a; b)$, и пишут

$$(a; b)-(c; d)=(a-c; b-d).$$

Деление. *Частным пар* $(a; b)$ и $(c; d)$ (пара $(c; d)$ не нулевая) называется пара $(x; y)$, удовлетворяющая равенству

$$(a; b)=(c; d)\cdot(x; y).$$

Из определения операции умножения пар и условия равенства пар следует, что частным пар $(a; b)$ и $(c; d)$ является пара

$$(x; y)=\left(\frac{ac+bd}{c^2+d^2}; \frac{bc-ad}{c^2+d^2}\right).$$

О паре $\left(\frac{ac+bd}{c^2+d^2}; \frac{bc-ad}{c^2+d^2}\right)$ говорят, что она получена в результате деления пары $(a; b)$ на пару $(c; d)$, и пишут

$$(a; b):(c; d)=\left(\frac{ac+bd}{c^2+d^2}; \frac{bc-ad}{c^2+d^2}\right).$$

Будем отождествлять пару вида $(a;\ 0)$ с действительным числом a. Нетрудно проверить, что при таком отождествлении суммой и произведением пар $(a;\ 0)$ и $(b;\ 0)$ будут пары

$$(a;\ 0)+(b;\ 0)=(a+b;\ 0),\quad (a;\ 0)\cdot(b;\ 0)=(a\cdot b;\ 0),$$

которые по принятому соглашению отождествляются с действительными числами $a+b$ и $a\cdot b$ соответственно.

Аналогично можно показать, что разность и частное пар $(a;\ 0)$ и $(b;\ 0)$ будут парами с нулевыми вторыми элементами, т. е. действительными числами (в случае частного предполагается, что $b\neq 0$).

Таким образом, множество всех действительных чисел является подмножеством построенного множества упорядоченных пар, причем все арифметические операции для действительных чисел, рассматриваемых как пары со вторым нулевым элементом, сохраняют тот же смысл, который они имели для действительных чисел.

Возьмем теперь пару $(0;\ 1)$ и по правилу умножения пар умножим ее на себя:

$$(0;\ 1)\cdot(0;\ 1)=(-1;\ 0). \tag{1}$$

В результате умножения получилась пара $(-1;\ 0)$, которая соответствует действительному числу -1. Введем для пары $(0;\ 1)$ специальное обозначение

$$(0;\ 1)=i.$$

Теперь равенство (1) можно записать в виде *)

$$i\cdot i=-1 \quad \text{или} \quad i^2=-1.$$

Простая проверка показывает, что любая пара действительных чисел $(a;\ b)$ может быть записана в виде

$$(a;\ b)=(a;\ 0)+(b;\ 0)\cdot(0;\ 1)$$

или (по принятому соглашению об отождествлении множества всех пар вида $(a;\ 0)$ с множеством всех действительных чисел и принятому обозначению для пары $(0;\ 1)$) — в виде

$$a+b\cdot i \quad \text{или} \quad a+bi. \tag{2}$$

При аксиоматическом способе введения комплексных чисел (см. п. 1.1) комплексное число обозначалось единым символом $a+b\cdot i$, в котором символам арифметических операций «$+$» и «$\cdot$» не приписывался какой-либо конкретный смысл. Теперь, когда введены операции сложения и умножения упорядоченных пар действительных чисел и показано, что любое комплексное число может быть записано в виде комбинации упорядоченных пар действительных чисел, символам этих арифметических операций в записи (2) можно придать смысл операций сложения и умножения пар.

§ 2. Геометрическое изображение и тригонометрическая форма записи комплексных чисел

2.1. Геометрическое изображение комплексного числа. Подобно тому как действительные числа можно изображать точками числовой прямой, комплексные числа можно изображать точками плоскости.

*) По сути дела, это равенство означает, что i является корнем уравнения

$$x^2+1=0.$$

Возможность такого изображения основана на отождествлении множества комплексных чисел $a+bi$ с множеством пар действительных чисел $(a; b)$, которые можно трактовать как координаты точек плоскости в прямоугольной системе координат Oxy.

Далее, с каждой точкой A координатной плоскости Oxy можно связать вектор $\overrightarrow{OA}$, выходящий из начала координат и оканчивающийся в точке A. Поэтому комплексные числа допускают и еще одну геометрическую интерпретацию: комплексное число $a+bi$ можно геометрически интерпретировать как вектор $\overrightarrow{OA}$ с координатами $(a; b)$ (рис. 3.1). Координаты вектора $\overrightarrow{OA}$ при этом будут такими же, как и координаты точки A, а именно $(a; b)$.

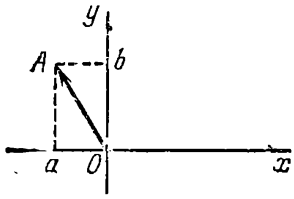


Рис. 3.1

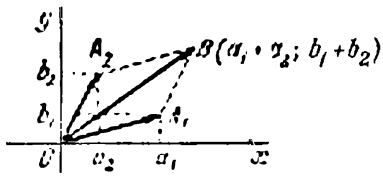


Рис. 3.2

2.2. Геометрическое изображение суммы и разности комплексных чисел. Геометрическая интерпретация комплексных чисел позволяет наглядно представить сумму и разность комплексных чисел. Пусть даны два комплексных числа $z_1=a_1+b_1i$ и $z_2=a_2+b_2i$. Их суммой будет комплексное число $z_1+z_2=(a_1+a_2)+(b_1+b_2)\,i$. В то же

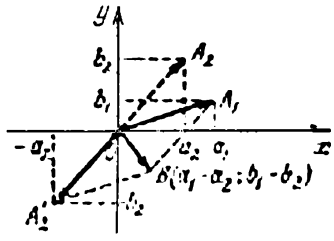


Рис. 3.3

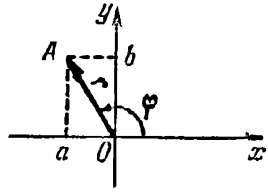


Рис. 3.4

время известно, что при сложении векторов их соответственные координаты складываются. Поэтому если вектор $\overrightarrow{OA_1}$ имеет координаты $(a_1; b_1)$ (рис. 3.2), а вектор $\overrightarrow{OA_2}$ — координаты $(a_2; b_2)$, то их сумма (вектор $\overrightarrow{OB}$) будет иметь координаты $(a_1+a_2;\ b_1+b_2)$. Вектор $\overrightarrow{OB}$ и есть геометрическое изображение суммы комплексных чисел z_1 и z_2.

Так как разность двух комплексных чисел $z_1=a_1+b_1i$ и $z_2=a_2+b_2i$ есть сумма комплексного числа z_1 и числа, противоположного комплексному числу z_2, то геометрически ее можно изобразить как сумму вектора $\overrightarrow{OA_1}$ с координатами $(a_1; b_1)$ и вектора $\overrightarrow{OA'_2}$ с координатами $(-a_2;\ -b_2)$ (рис. 3.3), т. е. как вектор $\overrightarrow{OB}$ с координатами $(a_1-a_2;\ b_1-b_2)$.

2.3. Тригонометрическая форма записи комплексного числа. Наряду с алгебраической формой записи комплексного числа в виде $a+bi$, употребляется и другая, тригонометрическая. Пусть комплексное число $z=a+bi\neq 0$ изображается вектором $\overrightarrow{OA}$ с координатами $(a;\ b)$ (рис. 3.4). Обозначим длину вектора $\overrightarrow{OA}$ буквой r:

$$r=|\overrightarrow{OA}|,$$

а угол, который вектор образует с положительным направлением оси Ox, — через φ (угол φ считаем измеренным в радианах).

По определению функций $\sin\varphi=b/r$, $\cos\varphi=a/r$, комплексное число $z=a+bi$ можно записать в виде

$$z=r\cdot(\cos\varphi+i\sin\varphi), \tag{1}$$

где $r=\sqrt{a^2+b^2}$, а угол φ определяется из условий

$$\sin\varphi=\frac{b}{\sqrt{a^2+b^2}},\quad \cos\varphi=\frac{a}{\sqrt{a^2+b^2}}. \tag{2}$$

Запись комплексного числа в виде (1) называется *тригонометрической формой* записи комплексного числа. Действительное число r называется *модулем* комплексного числа и обозначается $|z|$, а угол φ (измеренный в радианах) — *аргументом* комплексного числа z. Аргумент φ комплексного числа z обозначается $\operatorname{Arg} z$.

Если комплексное число не равно нулю, то модуль его положителен; если же $z=0$, т. е. $a=b=0$, то и модуль его равен нулю. Модуль любого комплексного числа определен однозначно.

Если комплексное число $z=a+bi$ не равно нулю, то его аргумент определяется формулами (2) с точностью до угла, кратного 2π. Если же $z=0$, то $r=0$ и аргумент комплексного числа, равного нулю, не определен.

Обычно для того, чтобы исключить неоднозначность, возникающую при вычислении аргумента комплексного числа, используют понятие *главного значения* аргумента комплексного числа (обозначение $\arg z$), считая, что $\arg z\in(-\pi;\ \pi]$. Аргумент комплексного числа z связан с главным значением аргумента соотношением

$$\operatorname{Arg} z=\arg z+2\pi k,\quad k\in\mathbf{Z}.$$

Пусть z_1 и z_2 — два отличных от нуля комплексных числа, записанных в тригонометрической форме:

$$z_1=r_1\cdot(\cos\varphi_1+i\sin\varphi_1),\quad z_2=r_2\cdot(\cos\varphi_2+i\sin\varphi_2).$$

Произведение двух комплексных чисел z_1 и z_2 есть комплексное число, модуль которого равен произведению модулей сомножителей, а аргумент — сумме аргументов сомножителей:

$$z_1\cdot z_2=r_1\cdot r_2\cdot[\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2)].$$

Вектор, изображающий произведение комплексных чисел z_1 и z_2, получается поворотом вектора z_1 против часовой стрелки на угол, равный φ_2, и растяжением его в $|z_2|$ раз (для случая $|z_2|>1$ см. рис. 3.5).

Частное двух комплексных чисел z_1 и z_2 есть комплексное число, модуль которого равен частному модулей делимого и дели-

теля, а аргумент частного двух отличных от нуля комплексных чисел равен разности аргументов делимого и делителя:

$$\frac{z_1}{z_2}=\frac{r_1}{r_2}\cdot[\cos(\varphi_1-\varphi_2)+i\sin(\varphi_1-\varphi_2)].$$

Вектор, изображающий частное двух комплексных чисел z_1 и z_2, получается поворотом вектора, изображающего комплексное число z_1,

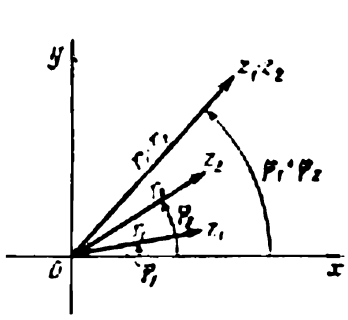


Рис. 3.5

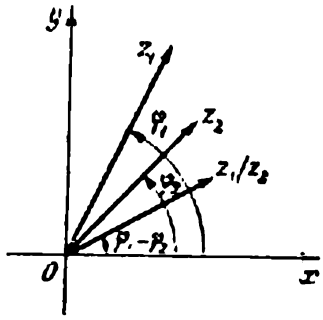


Рис. 3.6

по часовой стрелке на угол, равный φ_2, и сжатием его в $|z_2|$ раз (для случая $|z_2|>1$ см. рис. 3.6).

§ 3. Степень комплексного числа

3.1. Натуральная степень комплексного числа. Закон ассоциативности умножения комплексных чисел позволяет ввести понятие натуральной степени комплексного числа: *n*-й натуральной *степенью* комплексного числа z называется комплексное число w, полученное в результате умножения числа z на себя n раз:

$$w=\underbrace{z\cdot z\cdot\ldots\cdot z}_{n\text{ сомножителей}};$$

n-ю степень числа z обозначают z^n и пишут

$$w=z^n.$$

Число z называется *основанием степени*, а натуральное число n — *показателем степени*.

n-я степень комплексного числа, записанного в тригонометрической форме

$$z=r(\cos\varphi+i\sin\varphi),$$

вычисляется по *формуле Муавра*:

$$z^n=r^n(\cos n\varphi+i\sin n\varphi).$$

3.2. Корень *n*-й степени из комплексного числа. Под *корнем n-й степени* из комплексного числа z понимается множество комплексных чисел w, являющихся решениями уравнения

$$w^n=z. \qquad (1)$$

Корень n-й степени из комплексного числа z обозначается символом $\sqrt[n]{z}$.

Все корни n-й степени из комплексного числа z, заданного в тригонометрической форме

$$z = r(\cos\varphi + i\sin\varphi),$$

вычисляются по формуле

$$\sqrt[n]{z} = \sqrt[n]{r}\left(\cos\frac{\varphi+2\pi k}{n} + i\sin\frac{\varphi+2\pi k}{n}\right),$$

где $k = 0, 1, 2, \ldots, n-1$.

Геометрически все корни n-й степени из комплексного числа $z = r(\cos\varphi + i\sin\varphi)$ изображаются точками, лежащими на окружности с центром в начале координат, радиус которой равен $\sqrt[n]{r}$, а центральные углы между радиусами, проведенными в соседние точки, равны $2\pi/n$.

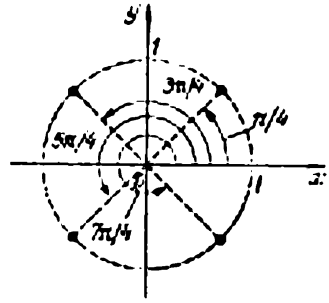


Рис. 3.7

Пример. Вычислить корни четвёртой степени из числа -1.

Число -1 в тригонометрической форме может быть записано так:

$$-1 = 1\cdot(\cos\pi + i\sin\pi).$$

Корни четвертой степени из числа -1 — это комплексные числа

$$\sqrt[4]{-1} = \sqrt[4]{1}\left(\cos\frac{\pi+2\pi k}{4} + i\sin\frac{\pi+2\pi k}{4}\right),$$

где $k = 0, 1, 2, 3$, т. е. комплексные числа (см. рис. 3.7)

$$\cos\frac{\pi}{4} + i\sin\frac{\pi}{4} = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i,$$

$$\cos\frac{3\pi}{4} + i\sin\frac{3\pi}{4} = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i,$$

$$\cos\frac{5\pi}{4} + i\sin\frac{5\pi}{4} = -\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i,$$

$$\cos\frac{7\pi}{4} + i\sin\frac{7\pi}{4} = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i.$$

Аналогичным образом в множестве комплексных чисел можно вычислить корень n-й степени из любого действительного числа. При этом хотя бы один корень из положительного действительного числа будет действительным.

Глава 4

АЛГЕБРА

Термин «алгебра» происходит от названия сочинения Мухаммеда аль-Хорезми «Альджебр аль-мукабала» (IX век), содержащего общие методы решения задач, сводящихся к уравнениям первой и второй степеней. К середине XVII века в основном сложилась современная алгебраическая символика. Вплоть до XVIII века под алгеброй понималась наука о буквенных вычислениях — тождественные преобразования буквенных формул *), нахождение формул для корней уравнений первой — четвертой степеней, логарифмы, прогрессии, комбинаторика.

В XVIII—XIX веках предмет алгебры — это прежде всего изучение многочленов, теория алгебраических уравнений с одним неизвестным, теория систем линейных уравнений с несколькими неизвестными, а также теория матриц и определителей.

Третий (современный) этап развития алгебры как науки об алгебраических операциях начался в середине XIX века и был связан с появлением примеров алгебраических операций над математическими объектами совсем иной природы, нежели действительные числа. Первыми такими примерами явились умножения подстановок и операции над комплексными числами.

§ 1. Многочлены одного переменного

1.1. Понятие многочлена. Арифметические операции над многочленами. Произведение вида ax^n, где a — отличный от нуля числовой коэффициент, а n — неотрицательное целое число, называется *одночленом* относительно переменной x. Число n называют *степенью* одночлена.

Два одночлена ax^n и bx^k называются *подобными*, если $n=k$. Замена суммы двух подобных одночленов ax^n+bx^n на одночлен $(a+b)x^n$ называется *приведением подобных членов*.

Многочленом (*полиномом*) n-й степени относительно переменной величины x называется выражение вида

$$P(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n, \qquad (1)$$

где n — неотрицательное целое число; a_0, a_1, a_2, $\ldots$, a_{n-1}, a_n — коэффициенты многочлена, причем коэффициент a_0, называемый старшим коэффициентом, считается не равным нулю.

Многочлен первой степени называют также *линейным* многочленом, многочлен второй степени — *квадратным*, а многочлен третьей степени — *кубичным* многочленом.

*) В алгебре, в отличие от арифметики, имеют дело не с числами, а с буквенными выражениями, называя их *алгебраическими выражениями*.

Запись многочлена в виде (1) называется записью по *убывающим степеням* переменной x; эта запись и будет использоваться в дальнейшем (если не будет оговорено противное). Каждое слагаемое в выражении (1) называется *членом* многочлена.

Из определения многочлена n-й степени следует, что существуют многочлены n-й степени для любого неотрицательного n. В частности, любое отличное от нуля действительное число является многочленом нулевой степени. Число 0 также считается многочленом; это единственный многочлен, степень которого не определена.

Два многочлена $P(x)$ и $Q(x)$ считают *равными* (или тождественно равными), если равны их коэффициенты при одинаковых степенях переменной x, и пишут

$$P(x) = Q(x).$$

Суммой многочленов

$$P(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n,$$
$$Q(x) = b_0x^m + b_1x^{m-1} + b_2x^{m-2} + \ldots + b_{m-1}x + b_m$$

называют многочлен $S(x)$, коэффициенты которого при каждой степени x равны сумме коэффициентов при этой степени x многочленов $P(x)$ и $Q(x)$. О многочлене $S(x)$ говорят, что он получен в результате сложения многочленов $P(x)$ и $Q(x)$, и пишут

$$S(x) = P(x) + Q(x).$$

Пример 1. Суммой многочленов $P(x) = x^4 - x^3 + 3$ и $Q(x) = 2x^3 - x + 2$ является многочлен $S(x) = x^4 + x^3 - x + 5$.

Произведением многочленов

$$P(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n,$$
$$Q(x) = b_0x^m + b_1x^{m-1} + b_2x^{m-2} + \ldots + b_{m-1}x + b_m$$

называют многочлен $M(x)$ степени $n+m$, коэффициенты которого $c_0, c_1, c_2, \ldots, c_{n+m-1}, c_{n+m}$ вычисляются по формулам

$$\begin{aligned}
c_0 &= a_0b_0,\\
c_1 &= a_1b_0 + a_0b_1,\\
&\ldots\ldots\ldots\\
c_k &= a_kb_0 + a_{k-1}b_1 + \ldots + a_1b_{k-1} + a_0b_k,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
c_{n+m} &= a_nb_m,
\end{aligned}$$

т. е. коэффициент c_i есть сумма всевозможных произведений коэффициентов a_l и b_k, для которых $l + k = i$. О многочлене $M(x)$ говорят, что он получен в результате умножения многочлена $P(x)$ на многочлен $Q(x)$, и пишут

$$M(x) = P(x) \cdot Q(x).$$

Пример 2. Произведением многочлена второй степени $P(x) = x^2 + 1$ на многочлен четвертой степени $Q(x) = 2x^4 - x + 3$ является многочлен шестой степени

$$M(x) = P(x)\,Q(x) = 2x^6 + 2x^4 - x^3 + 3x^2 - x + 3.$$

Операции сложения и умножения многочленов ассоциативны, коммутативны и связаны между собой законом дистрибутивности.

Противоположным для многочлена

$$P(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n$$

называют многочлен

$$-a_0x^n - a_1x^{n-1} - a_2x^{n-2} - \ldots - a_{n-1}x - a_n.$$

Многочлен, противоположный многочлену $P(x)$, обозначают $-P(x)$. Сумма многочлена $P(x)$ и противоположного ему многочлена $-P(x)$ равна нулю:

$$P(x) + (-P(x)) = 0.$$

Разностью многочленов $P(x)$ и $Q(x)$ называют многочлен $L(x)$ — сумму многочлена $P(x)$ и многочлена, противоположного многочлену $Q(x)$:

$$L(x) = P(x) + (-Q(x)).$$

О многочлене $L(x)$ говорят, что он получен в результате вычитания многочлена $Q(x)$ из многочлена $P(x)$, и пишут

$$L(x) = P(x) - Q(x).$$

1.2. Делители многочлена. Если для заданных многочленов $P(x)$ и $Q(x)$ найдется многочлен $G(x)$, удовлетворяющий равенству

$$P(x) = Q(x) \cdot G(x),$$

то говорят, что многочлен $P(x)$ *делится* (или нацело делится) на многочлен $Q(x)$; при этом многочлен $Q(x)$ называют *целым делителем* (или просто *делителем*) многочлена $P(x)$, а многочлен $G(x)$ — *частным* многочленов $P(x)$ и $Q(x)$. Если многочлен $Q(x)$ — делитель многочлена $P(x)$, то и частное $G(x)$ — делитель многочлена $P(x)$.

Основные свойства делителей многочлена.

1. Если многочлен $P(x)$ делится на многочлен $K(x)$, а $K(x)$ делится на многочлен $Q(x)$, то и $P(x)$ делится на $Q(x)$.

2. Если многочлены $P(x)$ и $K(x)$ делятся на многочлен $Q(x)$, то их сумма и разность также делятся на $Q(x)$.

3. Если многочлен $P(x)$ делится на многочлен $Q(x)$, то произведение $P(x)$ на любой многочлен $K(x)$ также делится на $Q(x)$.

4. Всякий многочлен $P(x)$ делится на любой многочлен нулевой степени, т. е. на число, отличное от нуля.

5. Многочлены $cP(x)$ $(c \neq 0)$, и только они, являются делителями многочлена $P(x)$, имеющими такую же степень, что и $P(x)$.

6. Многочлены $P(x)$ и $Q(x)$ тогда и только тогда одновременно делятся друг на друга, когда $P(x) = cQ(x)$ $(c \neq 0)$.

Наибольший общий делитель двух многочленов. Многочлен $R(x)$ называется *общим делителем* для многочленов $P(x)$ и $Q(x)$, если он служит делителем для каждого из этих многочленов. Согласно свойству 4) (см. выше) общими делителями произвольных многочленов $P(x)$ и $Q(x)$ будут все многочлены нулевой степени (т. е. все действительные числа, за исключением числа 0). Если других общих делителей два многочлена не имеют, то эти многочлены называются *взаимно простыми*.

Наибольшим общим делителем многочленов $P(x)$ и $Q(x)$ называется многочлен $D(x)$, который является их общим делителем и вместе с тем сам делится на любой общий делитель этих многочленов.

Если $D(x)$ — наибольший общий делитель многочленов $P(x)$ и $Q(x)$, то наибольшим общим делителем этих многочленов будет и многочлен $cD(x)$, где c — произвольное число, отличное от нуля, т. е. наибольший общий делитель двух многочленов определен с точностью до постоянного множителя. Чтобы добиться полной однозначности в определении наибольшего общего делителя двух многочленов, обычно

выставляют следующее условие: из всех многочленов $cD(x)$ наибольшим общим делителем считается тот, у которого старший коэффициент равен единице.

Теперь можно сказать, что два многочлена взаимно простые, если их наибольший делитель равен единице.

1.3. Деление многочленов. Схема Горнера. Теорема Безу. Пусть $P(x)$ и $Q(x)$ — заданные многочлены, и пусть степень многочлена $P(x)$ больше или равна степени многочлена $Q(x)$. Если многочлен $P(x)$ не делится (нацело) на многочлен $Q(x)$, т. е. не существует многочлена $G(x)$, удовлетворяющего равенству

$$P(x) = Q(x) \cdot G(x),$$

то вводят операцию *деления многочленов с остатком*.

Разделить многочлен $P(x)$ на многочлен $Q(x)$ с остатком — это значит найти два многочлена $G(x)$ и $R(x)$, удовлетворяющие равенству

$$P(x) = Q(x) \cdot G(x) + R(x), \tag{2}$$

причем степень многочлена $R(x)$ меньше степени многочлена $Q(x)$.

Многочлен $P(x)$ называется *делимым*, $Q(x)$ — *делителем*, $G(x)$ — *частным*, а $R(x)$ — *остатком*. При делении многочлена $P(x)$ на многочлен $Q(x)$ многочлены $G(x)$ и $R(x)$ находятся однозначно.

Для деления многочлена на многочлен обычно применяется правило деления «углом». Для этого располагают многочлены $P(x)$ и $Q(x)$ по убывающим степеням x и находят старший член частного $G(x)$ из условия, что при умножении его на старший член делителя $Q(x)$ должен получиться старший член делимого $P(x)$. Найденный старший член частного умножают затем на делитель и полученный многочлен вычитают из делимого. В результате вычитания получается многочлен $P_1(x)$. Следующим членом частного будет такой одночлен, который, будучи умноженным на старший член делителя $Q(x)$, дает старший член многочлена $P_1(x)$, и т. д. Процесс продолжается до тех пор, пока степень очередной разности (т. е. степень некоторого многочлена $P_k(x)$) не станет меньше степени делителя. Многочлен $P_k(x)$ и будет остатком от деления многочлена $P(x)$ на многочлен $Q(x)$.

В результате деления многочлена n-й степени $P(x)$ на многочлен k-й степени $Q(x)$ $(k \leqslant n)$ в частном получается многочлен степени $n-k$.

Пример. Разделить многочлен $P(x) = x^4 + x^3 + 3x^2 + 3x + 2$ на многочлен $Q(x) = x^2 + x + 1$.

Старший член многочлена $P(x)$ есть x^4, а старший член многочлена $Q(x)$ — член x^2. Разделив x^4 на x^2, получим x^2. Одночлен x^2 — старший член частного. Умножая одночлен x^2 на многочлен $Q(x)$, получаем многочлен $x^4 + x^3 + x^2$. Вычтем его из многочлена $P(x)$:

$$(x^4 + x^3 + 3x^2 + 3x + 2) - (x^4 + x^3 + x^2) = 2x^2 + 3x + 2 = P_1(x).$$

Степень многочлена $P_1(x)$ не меньше степени многочлена $Q(x)$, поэтому можно продолжить процесс деления; делим старший член $2x^2$ многочлена $P_1(x)$ на старший член многочлена $Q(x)$:

$$2x^2/x^2 = 2.$$

Многочлен нулевой степени (число 2) будет следующим членом частного.

Умножаем $Q(x)$ на число 2; получаем многочлен $2x^2 + 2x + 2$, который вычитаем из $P_1(x)$:

$$P_1(x) - (2x^2 + 2x + 2) = 2x^2 + 3x + 2 - (2x^2 + 2x + 2) = x = P_2(x).$$

На этом процесс деления прекращается, так как степень многочлена $P_2(x)$ меньше степени многочлена $Q(x)$. Многочлен $P_2(x)$ и будет остатком от деления $P(x)$ на $Q(x)$.

Вся описанная процедура деления многочлена $P(x)$ на многочлен $Q(x)$ коротко отражается в следующей записи:

$$\begin{array}{r|l} x^4+x^3+3x^2+3x+2 & x^2+x+1 \\ \underline{-\,x^4+x^3+x^2} & \\ 2x^2+3x+2 & x^2+2 \\ \underline{-\,2x^2+2x+2} & \\ x & \end{array}$$

Итак, частное от деления многочлена $P(x)=x^4+x^3+3x^2+3x+2$ на многочлен $Q(x)=x^2+x+1$ есть многочлен $G(x)=x^2+2$, а остаток — многочлен $R(x)=x$, т. е. многочлен $P(x)$ может быть записан в виде

$$x^4+x^3+3x^2+3x+2=(x^2+x+1)(x^2+2)+x.$$

Схема Горнера. Теорема Безу. Для деления многочлена n-й степени

$$P(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$$

на многочлен первой степени $x-c$ обычно используется метод сокращенного деления (называемый *схемой Горнера*). В соответствии с определением операции деления многочленов с остатком (см. формулу (2) предыдущего пункта) в результате деления многочлена n-й степени на линейный многочлен $x-c$ в остатке может получиться либо многочлен нулевой степени (т. е. отличное от нуля число), либо нуль, а степень частного равна $n-1$.

Пусть частное многочленов $P(x)$ и $x-c$ имеет вид

$$G(x)=\alpha_0x^{n-1}+\alpha_1x^{n-2}+\ldots+\alpha_{n-2}x+\alpha_{n-1};$$

а остаток $R(x)$ равен числу β. Для рассматриваемого случая формула (2) принимает вид

$$a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n=$$
$$=(x-c)(\alpha_0x^{n-1}+\alpha_1x^{n-2}+\ldots+\alpha_{n-2}x+\alpha_{n-1})+\beta. \qquad (3)$$

Раскрывая скобки и приводя подобные члены в правой части равенства (3), по определению равенства многочленов получаем систему линейных уравнений для нахождения α_0, α_1, α_2, ..., α_{n-2}, α_{n-1}, β:

$$\begin{aligned} \alpha_0 &= a_0, \\ \alpha_1 &= a_1+c\alpha_0, \\ \alpha_2 &= a_2+c\alpha_1, \\ &\cdots\cdots\cdots \\ \alpha_{n-1} &= a_{n-1}+c\alpha_{n-2}, \\ \beta &= a_n+c\alpha_{n-1}. \end{aligned}$$

Эту систему называют *схемой Горнера*.

Первое уравнение дает значение $\alpha_0=a_0$. Подставляя это значение α_0 во второе уравнение, получаем $\alpha_1=a_1+ca_0$. Подставляя полученное значение α_1 в третье уравнение, получаем значение α_2 и т. д. Последним будет найдено выражение для остатка β:

$$\beta=a_0c^n+a_1c^{n-1}+a_2c^{n-2}+\ldots+a_{n-1}c+a_n.$$

Это равенство известно под названием теоремы Безу.

Многочлен $a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$ при делении на $x-c$ дает остаток, равный значению этого многочлена при $x=c$.

1.4. Алгоритм Евклида нахождения наибольшего общего делителя двух многочленов. Пусть $P(x)$ и $Q(x)$ — два произвольных многочлена и степень многочлена $P(x)$ больше или равна степени многочлена $Q(x)$. Делим многочлен $P(x)$ на многочлен $Q(x)$, т. е. находим частное $G(x)$ и остаток $R(x)$ от деления этих многочленов. Делим затем $Q(x)$ на $R(x)$ и получаем остаток $R_1(x)$; делим $R(x)$ на $R_1(x)$ и получаем остаток $R_2(x)$; делим $R_1(x)$ на $R_2(x)$ и т. д. Поскольку степени остатков все время понижаются, то в этой цепочке последовательных делений через конечное число шагов появятся такие два многочлена $R_{n-1}(x)$ и $R_n(x)$, что $R_{n-1}(x)$ разделится на $R_n(x)$ нацело. Остаток $R_n(x)$, на который нацело делится предыдущий остаток $R_{n-1}(x)$, и будет наибольшим общим делителем многочленов $P(x)$ и $Q(x)$. При этом, правда, старший коэффициент этого наибольшего общего делителя не обязательно будет равен единице.

Наибольший общий делитель двух многочленов со старшим коэффициентом, равным единице, получится, если все коэффициенты полученного общего делителя разделить на коэффициент при старшем члене.

Для наглядности запишем алгоритм Евклида в виде цепочки равенств;

$$\begin{aligned}
P(x) &= Q(x)\cdot G(x)+R(x),\\
Q(x) &= R(x)\cdot G_1(x)+R_1(x),\\
R(x) &= R_1(x)\cdot G_2(x)+R_2(x),\\
R_1(x) &= R_2(x)\cdot G_3(x)+R_3(x),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
R_{n-2}(x) &= R_{n-1}(x)\cdot G_n(x)+R_n(x),\\
R_{n-1}(x) &= R_n(x)\cdot G_{n+1}(x).
\end{aligned}$$

Отметим еще раз, что наибольшим общим делителем многочленов $P(x)$ и $Q(x)$ будет многочлен $R_n(x)$, разделенный на старший коэффициент.

Алгоритм Евклида нахождения наибольшего общего делителя двух многочленов полностью аналогичен алгоритму Евклида нахождения наибольшего общего делителя двух натуральных чисел.

Пример. Найти наибольший общий делитель многочленов

$$P(x)=x^5-x^4-2x^3+2x^2+x-1,\qquad Q(x)=x^3-1.$$

Разделив «углом» многочлен $P(x)$ на многочлен $Q(x)$;

$$\begin{array}{r|l}
-\begin{array}{l} x^5-x^4-2x^3+2x^2+x-1\\ x^5\qquad\qquad -x^2\end{array} & \dfrac{x^3-1}{x^2-x-2}\\
\hline
-\begin{array}{l} -x^4-2x^3+3x^2+x\\ -x^4\qquad\qquad +x\end{array} & \\
\hline
-\begin{array}{l} -2x^3+3x^2\quad -1\\ -2x^3\qquad\quad +2\end{array} & \\
\hline
3x^2\quad -3 &
\end{array}$$

получим частное — многочлен $G(x)=x^2-x-2$ и остаток — многочлен $R(x)=3x^2-3$.

Делим многочлен $Q(x)$ на остаток $R(x)$:

$$\begin{array}{r|l}
-\begin{array}{l} x^3-1\\ x^3-x\end{array} & \dfrac{3x^2-3}{\frac{1}{3}x}\\
\hline
x-1 &
\end{array}$$

Получим частное $G_1(x)=\frac{1}{3}x$ и остаток $R_1(x)=x-1$. По алгоритму Евклида

делим остаток $R(x)$ на остаток $R_1(x)$:

$$\begin{array}{r|l} 3x^2-3 & x-1 \\ \underline{-\,3x^2-3x} & \overline{3x+3} \\ \quad 3x-3 & \\ \underline{-\quad 3x-3} & \\ \quad 0 & \end{array}$$

Многочлен $R(x)=3x^2-3$ нацело разделился на многочлен $R_1(x)=x-1$. Следовательно, многочлен $x-1$ — наибольший общий делитель многочленов $P(x)$ и $Q(x)$.

1.5. Корни многочлена. Пусть

$$P(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n$$

— некоторый многочлен n-й степени. Возьмем число c и подставим его вместо переменной x в многочлен $P(x)$. Полученное число $a_0c^n+a_1c^{n-1}+a_2c^{n-2}+\ldots+a_{n-1}c+a_n$ называется *значением многочлена* $P(x)$ при $x=c$ и обозначается $P(c)$. Если $P(c)=0$, то число c называется *корнем многочлена* $P(x)$.

Связь между корнем многочлена и существованием делителя многочлена устанавливается следующей теоремой.

Если число c — корень многочлена n-й степени $P(x)$, то многочлен $P(x)$ делится на линейный многочлен $x-c$, т. е. многочлен $P(x)$ можно представить в виде

$$P(x)=(x-c)\,Q(x),$$

где $Q(x)$ — многочлен степени $n-1$.

Может оказаться, что многочлен n-й степени $P(x)$ делится не только на линейный многочлен $x-c$, но и на его более высокую степень, т. е. на многочлен вида $(x-c)^k$, где k — натуральное число, большее единицы. Число c называется *корнем кратности* k многочлена $P(x)$ (или *k-кратным корнем*), если многочлен $P(x)$ делится на многочлен $(x-c)^k$, но не делится на многочлен $(x-c)^{k+1}$. Если многочлен $P(x)$ делится на линейный многочлен $x-c$, но не делится на $(x-c)^2$, то число c называют *простым корнем* многочлена $P(x)$.

Для того чтобы число c было k-кратным корнем многочлена $P(x)$, необходимо и достаточно, чтобы число c было $(k-1)$-кратным корнем производной многочлена $P(x)$, т. е. многочлена $P'(x)$.

1.6. Формулы сокращенного умножения. Поскольку число c — корень многочлена x^n-c^n (n любое), а число $(-c)$ — корень многочленов x^n-c^n (n четное) и x^n+c^n (n нечетное), то:

1) многочлен x^n-c^n делится на двучлен $x-c$ при любом натуральном n;

2) многочлен x^n-c^n делится на двучлен $x+c$ при любом четном n;

3) многочлен x^n+c^n делится на двучлен $x+c$ при любом нечетном n.

В результате деления получаются соответственно многочлены $(n-1)$-й степени:

$$\frac{x^n-c^n}{x-c}=x^{n-1}+x^{n-2}c+x^{n-3}c^2+\ldots+xc^{n-2}+c^{n-1}=$$

$$=\sum_{k=0}^{n-1}x^{n-k-1}c^k \qquad (n \text{ любое});$$

$$\frac{x^n - c^n}{x + c} = x^{n-1} - x^{n-2}c + x^{n-3}c^2 - \ldots + xc^{n-2} - c^{n-1} =$$

$$= \sum_{k=0}^{n-1} (-1)^k x^{n-k-1} c^k \qquad (n \text{ четное});$$

$$\frac{x^n + c^n}{x + c} = x^{n-1} - x^{n-2}c + x^{n-3}c^2 - \ldots - xc^{n-2} + c^{n-1} =$$

$$= \sum_{k=0}^{n-1} (-1)^k x^{n-k-1} c^k \qquad (n \text{ нечетное}).$$

В частности, из приведенных равенств следуют простейшие формулы, называемые *формулами сокращенного умножения*:

$$(x + c)(x - c) = x^2 - c^2,$$
$$(x + c)(x^2 - xc + c^2) = x^3 + c^3,$$
$$(x - c)(x^2 + xc + c^2) = x^3 - c^3.$$

1.7. Формулы Виета. Между корнями $c_1, c_2, \ldots, c_n$ многочлена n-й степени

$$P(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_{n-1} x + a_n$$

со старшим коэффициентом, равным единице (каждый корень взят столько раз, какова его кратность), существуют зависимости, называемые *формулами Виета*:

$$c_1 + c_2 + c_3 + \ldots + c_{n-1} + c_n = -a_1,$$
$$c_1 c_2 + c_1 c_3 + \ldots + c_{n-1} c_n = a_2,$$
$$c_1 c_2 c_3 + c_1 c_2 c_4 + \ldots + c_{n-2} c_{n-1} c_n = -a_3,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$c_1 c_2 c_3 \ldots c_{n-1} c_n = (-1)^n a_n.$$

Если $P(x)$ — квадратный трехчлен, т. е.

$$P(x) = x^2 + px + q,$$

а числа c_1, c_2 — его корни, то формулы Виета имеют вид

$$c_1 + c_2 = -p, \quad c_1 c_2 = q.$$

Если $P(x)$ — кубичный многочлен, т. е.

$$P(x) = x^3 + px^2 + qx + r,$$

а числа c_1, c_2, c_3 — его корни, то формулы Виета имеют вид

$$c_1 + c_2 + c_3 = -p, \quad c_1 c_2 + c_2 c_3 + c_1 c_3 = q, \quad c_1 c_2 c_3 = -r.$$

1.8. Основная теорема алгебры. Вводя понятие корней многочлена, мы не ставили вопроса о том, всякий ли многочлен с действительными коэффициентами имеет корни. Известно, что существуют многочлены с действительными коэффициентами, не имеющие действительных корней ($x^2 + 1$ — один из таких многочленов), и именно этот факт послужил одной из причин введения новых математических объектов — комплексных чисел. Можно было бы ожидать, что существуют многочлены, не имеющие корней даже в множестве комплексных чисел, особенно если рассматривать многочлены с любыми комплексными коэффициентами (частным случаем которых являются многочле-

ны с действительными коэффициентами). Если бы это было так, то система комплексных чисел нуждалась бы в дальнейшем расширении. На самом деле, однако, справедлива следующая основная теорема алгебры.

Всякий многочлен n-й степени с комплексными коэффициентами в множестве комплексных чисел имеет ровно n корней, если каждый кратный корень считать такое число раз, какова его кратность.

Все доказательства этой теоремы (а их найдено весьма много) основаны на использовании свойства непрерывности множеств действительных и комплексных чисел, вследствие чего все эти доказательства нельзя считать чисто алгебраическими.

Основная теорема алгебры справедлива и при $n=0$, так как многочлен нулевой степени корней не имеет. Основная теорема алгебры неприменима лишь к нулевому многочлену (числу нуль), степень которого не определена.

1.9. Разложение многочлена на множители. По основной теореме алгебры многочлен n-й степени

$$P(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n$$

с любыми комплексными (и, в частности, действительными) коэффициентами имеет n корней (с учетом их кратностей) и, следовательно, делится нацело на линейные многочлены

$$x-c_1,\quad x-c_2,\quad \ldots,\quad x-c_n,$$

где $c_1, c_2, \ldots, c_n$ — корни многочлена; значит, многочлен $P(x)$ представим в виде произведения n линейных сомножителей:

$$P(x)=a_0(x-c_1)(x-c_2)\ldots(x-c_n).$$

Разложение многочлена $P(x)$ в произведение линейных многочленов единственно с точностью до порядка следования сомножителей.

1.10. Некоторые следствия основной теоремы алгебры для многочленов с действительными коэффициентами. Если комплексное (но не действительное) число $c=\alpha+\beta i$ является корнем многочлена n-й степени с действительными коэффициентами, то и комплексно сопряженное число $\bar{c}=\alpha-\beta i$ будет корнем данного многочлена.

Из свойства сопряженности комплексных корней многочлена с действительными коэффициентами следует, что всякий многочлен нечетной степени с действительными коэффициентами имеет по крайней мере один действительный корень.

Пусть $P(x)$ — многочлен n-й степени с действительными коэффициентами, $c=\alpha+\beta i$ — его корень. Тогда $\bar{c}=\alpha-\beta i$ — также корень многочлена $P(x)$, и, следовательно, многочлен $P(x)$ может быть представлен в виде

$$P(x)=(x-c)(x-\bar{c})Q(x),$$

где $Q(x)$ — многочлен $(n-2)$-й степени. Произведение

$$(x-c)(x-\bar{c})=x^2-(c+\bar{c})x+c\bar{c}=x^2-2\alpha x+\alpha^2+\beta^2$$

есть многочлен второй степени с действительными коэффициентами, поэтому все коэффициенты многочлена $Q(x)$ — также действительные числа.

Итак, всякий многочлен с действительными коэффициентами можно представить, и притом единственным способом (с точностью до порядка сомножителей), в виде произведения старшего коэффи-

циента a_0, нескольких линейных многочленов с действительными коэффициентами вида $x-c$, соответствующих его действительным корням, и квадратных многочленов $x^2-(c+\bar{c})x+c\bar{c}$, соответствующих парам сопряженных комплексных корней.

Среди многочленов с действительными коэффициентами и со старшим коэффициентом, равным единице, неразложимыми на множители меньшей степени в множестве действительных чисел (*неприводимыми многочленами*) являются лишь линейные многочлены $x-c$ и квадратные многочлены вида $x^2-(c+\bar{c})x+c\bar{c}$.

Установленное свойство многочленов с действительными коэффициентами аналогично свойству разложимости любого натурального числа на простые сомножители. Для любого многочлена $P(x)$ n-й степени с действительными коэффициентами и первым коэффициентом, равным единице, можно написать разложение, аналогичное каноническому разложению натуральных чисел (см. п. 1.3, гл. 2):

$$P(x)=(x-c_1)^{k_1}(x-c_2)^{k_2}\ldots(x-c_l)^{k_l}[x^2-(c_{l+1}+\bar{c}_{l+1})x+ \\ +c_{l+1}\bar{c}_{l+1}]^{k_{l+1}}\ldots[x^2-(c_{l+m}+\bar{c}_{l+m})x+c_{l+m}\bar{c}_{l+m}]^{k_{l+m}}, \\ k_1+k_2+\ldots+k_l+2(k_{l+1}+\ldots+k_{l+m})=n,$$

где c_1, c_2, ..., c_l — действительные корни многочлена с кратностями k_1, k_2, ..., k_l, а $\{c_{l+1};\ \bar{c}_{l+1}\}$, ..., $\{c_{l+m};\ \bar{c}_{l+m}\}$ — пары комплексно сопряженных корней с кратностями k_{l+1}, ..., k_{l+m} соответственно.

§ 2. Многочлены нескольких переменных

2.1. Одночлены и многочлены нескольких переменных. Выражение вида

$$ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n},$$

где a — отличный от нуля числовой коэффициент, а k_1, k_2, ..., k_n — неотрицательные целые числа, называется *одночленом* от n переменных x_1, x_2, ..., x_n. Два одночлена

$$ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \quad \text{и} \quad bx_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}$$

называются *подобными*, если $k_1=l_1$, $k_2=l_2$, ..., $k_n=l_n$. Замена суммы двух подобных одночленов $ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}+bx_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$ на одночлен $(a+b)x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$ называется *приведением подобных членов*.

Многочленом $P(x_1, x_2, \ldots, x_n)$ от n переменных x_1, x_2, ..., x_n называется сумма конечного числа одночленов вида $ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$. Одночлены, образующие многочлен, называются *членами* многочлена.

При этом предполагается, что в многочлене $P(x_1, x_2, x_3, \ldots, x_n)$ все подобные члены уже приведены и что рассматриваются лишь члены с отличными от нуля коэффициентами. *Степенью многочлена* $P(x_1, x_2, x_3, \ldots, x_n)$ по отношению к некоторой выбранной переменной x_i $(i=1, 2, 3, \ldots, n)$ считается наивысший показатель, с которым переменная x_i входит в многочлен.

Степенью одночлена (по совокупности переменных)

$$x_1^{k_1}x_2^{k_2}x_3^{k_3}\ldots x_n^{k_n}$$

называют число

$$k_1 + k_2 + k_3 + \ldots + k_n,$$

т. е. сумму показателей степеней всех переменных x_i, входящих в данный одночлен. Наибольшее значение суммы $k_1 + k_2 + k_3 + \ldots + k_n$, встречающееся в каком-либо из членов многочлена, называется *степенью многочлена* по совокупности всех переменных.

Многочлен $P(x_1, x_2, x_3, \ldots, x_n)$ называют *однородным*, если степени всех одночленов, входящих в многочлен, равны.

2.2. Лексикографическое расположение членов многочлена. В то время как для многочленов от одной переменной существуют два естественных способа расположения членов многочлена — по убывающим и по возрастающим степеням переменной, для многочленов от нескольких переменных таких естественных способов уже не существует. Существует, однако, часто используемый способ вполне определенного расположения членов многочлена от нескольких переменных, называемый лексикографическим. Этот способ аналогичен способу расположения слов в словарях: считая буквы упорядоченными так, как это принято в алфавите, мы определяем взаимное расположение двух слов в словаре по их первым буквам; если же первые буквы совпадают, то по вторым буквам, и т. д.

Пусть $P(x_1, x_2, x_3, \ldots, x_n)$ — многочлен от n переменных $x_1, x_2, x_3, \ldots, x_n$, и пусть

$$x_1^{k_1} x_2^{k_2} x_3^{k_3} \ldots x_n^{k_n}, \quad x_1^{l_1} x_2^{l_2} x_3^{l_3} \ldots x_n^{l_n}$$

— два различных члена данного многочлена. Следовательно, хотя бы одна из разностей показателей степеней

$$k_1 - l_1, \quad k_2 - l_2, \ldots, k_n - l_n \tag{1}$$

отлична от нуля. Член $x_1^{k_1} x_2^{k_2} x_3^{k_3} \ldots x_n^{k_n}$ считается *выше* члена $x_1^{l_1} x_2^{l_2} x_3^{l_3} \ldots x_n^{l_n}$, если первая отличная от нуля разность в (1) положительна. В противном случае член $x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$ считается *ниже* члена $x_1^{l_1} x_2^{l_2} \ldots x_n^{l_n}$. Другими словами, член $x_1^{k_1} x_2^{k_2} x_3^{k_3} \ldots x_n^{k_n}$ выше члена $x_1^{l_1} x_2^{l_2} x_3^{l_3} \ldots x_n^{l_n}$, если $k_1 > l_1$ или если $k_1 = l_1$, то $k_2 > l_2$, и т. д. Заметим, что из того, что член $x_1^{k_1} x_2^{k_2} x_3^{k_3} \ldots x_n^{k_n}$ выше члена $x_1^{l_1} x_2^{l_2} x_3^{l_3} \ldots x_n^{l_n}$, вовсе не следует, что степень члена $x_1^{k_1} x_2^{k_2} x_3^{k_3} \ldots x_n^{k_n}$ больше степени члена $x_1^{l_1} x_2^{l_2} x_3^{l_3} \ldots x_n^{l_n}$. Так, член $x_1^3 x_2$ считается выше члена $x_1^2 x_2^5$, хотя его степень (по совокупности переменных) меньше.

Сравнивая между собой все члены многочлена $P(x_1, x_2, \ldots, x_n)$, выделяем из них тот, который выше всех остальных, и записываем его на первом месте. Сравниваем оставшиеся члены многочлена и выделяем из них наивысший, который записываем на втором месте, и т. д. В результате получим запись многочлена от n переменных, в которой каждый член, начиная со второго, ниже предыдущего.

Эта форма записи называется *лексикографическим способом* записи многочлена от нескольких переменных. Член многочлена, стоящий на первом месте, считается высшим членом многочлена.

Пример. Пусть

$$P(x_1, x_2, x_3) = x_1^2 x_2^2 x_3^3 + x_1^3 x_2^2 x_3 + x_1 x_2^2 x_3^3 + x_1 x_2^2 x_3 \tag{2}$$

многочлен трех переменных. Запишем многочлен, располагая его члены в лексикографическом порядке. В записи (2) второй член выше всех остальных, так как в этом члене переменная x_1 возводится в наибольшую степень по сравнению с остальными членами. Значит, член $x_1^3x_2^2x_3$ будет первым в лексикографической записи многочлена (2). Вторым членом в лексикографической записи будет первый член записи (2), так как в этом члене переменная x_1 возводится в наибольшую степень среди оставшихся членов. Из двух оставшихся членов (третьего и четвертого в записи (2)) третий будет выше четвертого, так как в этих членах степени переменных x_1 и x_2 равны, а переменная x_3 в третьем члене возводится в бо́льшую степень, нежели в четвертом.

Запись многочлена $P(x_1, x_2, x_3)$ с членами, расположенными в лексикографическом порядке, будет следующей:

$$P(x_1, x_2, x_3) = x_1^3x_2^2x_3 + x_1^2x_2^2x_3^5 + x_1x_2^2x_3^3 + x_1x_2^2x_3.$$

§ 3. Рациональные алгебраические дроби

3.1. Множество рациональных алгебраических дробей. Выражение вида

$$\frac{P}{Q} \quad (\text{или } P/Q), \tag{1}$$

где P и Q — многочлены одной или нескольких переменных, называют *алгебраической дробью*.

Многочлен P называется *числителем* дроби, а многочлен Q — ее *знаменателем*.

Рациональную алгебраическую дробь называют также *рациональным алгебраическим выражением*.

Рациональная алгебраическая дробь (1) определена при всех значениях переменной x, за исключением конечного набора значений, обращающих многочлен $Q(x)$ в нуль (корней многочлена $Q(x)$)*). Множество значений переменной x, при которых знаменатель дроби отличен от нуля, называют *множеством допустимых значений* алгебраической дроби.

Две рациональные дроби $P(x)/Q(x)$ и $M(x)/N(x)$ называются *равными*:

$$\frac{P(x)}{Q(x)} = \frac{M(x)}{N(x)},$$

если $P(x)\cdot N(x) = Q(x)\cdot M(x)$ при всех значениях x, при которых $Q(x) \neq 0$ и $N(x) \neq 0$.

Из определения равенства дробей следует, что если числитель и знаменатель дроби умножить на один и тот же многочлен $K(x)$, то получится дробь, равная исходной:

$$\frac{P(x)}{Q(x)} = \frac{P(x)\cdot K(x)}{Q(x)\cdot K(x)}.$$

---

*) Далее, если не оговорено противное, будем рассматривать дроби, числитель и знаменатель которых — многочлены одной переменной.

Сумма, произведение, разность и частное двух дробей P/Q и M/N вычисляются по правилам

$$\frac{P}{Q}+\frac{M}{N}=\frac{P\cdot N+Q\cdot M}{Q\cdot N};\quad \frac{P}{Q}\cdot\frac{M}{N}=\frac{P\cdot M}{Q\cdot N},$$

$$\frac{P}{Q}-\frac{M}{N}=\frac{P\cdot N-Q\cdot N}{Q\cdot N};\quad \frac{P}{Q}:\frac{M}{N}=\frac{P\cdot N}{Q\cdot M},$$

причем частное определено при всех тех допустимых значениях переменной x дробей P/Q и M/N, при которых $M(x)\neq 0$.

Операции сложения и умножения рациональных алгебраических дробей ассоциативны, коммутативны и связаны законом дистрибутивности умножения относительно сложения.

Рациональная алгебраическая дробь P/Q называется *несократимой*, если многочлены P и Q взаимно простые (см. п. 1.2). Всякая рациональная алгебраическая дробь равна некоторой несократимой дроби, определяемой однозначно с точностью до числового множителя, общего для числителя и знаменателя. Для того чтобы представить рациональную алгебраическую дробь в виде несократимой дроби, необходимо найти наибольший общий делитель многочленов P и Q и произвести сокращение дроби.

Более строго, множество рациональных алгебраических дробей может быть определено следующим образом.

Рациональными алгебраическими дробями называют упорядоченные пары $(P; Q)$ двух многочленов, определяемые перечисленными ниже свойствами.

Две упорядоченные пары (рациональные дроби) $(P; Q)$ и $(M; N)$ называют *равными* и пишут

$$(P; Q)=(M; N),$$

если $P\cdot N=Q\cdot M$ при всех значениях, при которых $Q(x)\neq 0$ и $N(x)\neq 0$.

Суммой двух пар $(P; Q)$ и $(M; N)$ называют пару $(P\cdot N+Q\cdot M; Q\cdot N)$. Говорят, что она получена в результате сложения пар $(P; Q)$, $(M; N)$, и пишут

$$(P; Q)+(M; N)=(P\cdot N+Q\cdot M; Q\cdot N).$$

Ассоциативность и коммутативность сложения пар следует из определения суммы пар и соответствующих свойств сложения и умножения многочленов.

Разностью пар $(P; Q)$ и $(M; N)$ называют пару $(X; Y)$, удовлетворяющую равенству

$$(M; N)+(X; Y)=(P; Q).$$

Из определения суммы пар и определения равных пар следует, что

$$(X; Y)=(P\cdot N-Q\cdot M; Q\cdot N).$$

О паре $(P\cdot N-Q\cdot M; Q\cdot N)$ говорят, что она получена в результате вычитания пары $(M; N)$ из пары $(P; Q)$, и пишут

$$(P; Q)-(M; N)=(P\cdot N-Q\cdot M; Q\cdot N).$$

Из определения разности пар следует, что

$$(P; Q)-(P; Q)=(0; Q).$$

Пары вида $(0; Q)$ называют *нулевыми парами*.

Произведением двух пар $(P; Q)$ и $(M; N)$ называют пару $(P\cdot M; Q\cdot N)$. Говорят, что она получена в результате умножения пар $(P; Q)$ и $(M; N)$, и пишут

$$(P; Q)\cdot(M; N)=(P\cdot M; Q\cdot N).$$

Из определения произведений пар и свойств арифметических операций над многочленами следует, что операция умножения пар ассоциативна и коммутативна, а также дистрибутивна по отношению к сложению.

В множестве упорядоченных пар многочленов обратная умножению операция деления пары $(P; Q)$ на ненулевую пару $(M; N)$ выполнима при всех значениях x, при которых многочлен $M(x) \neq 0$.

Частным пар $(P; Q)$ и $(M; N)$ называют пару $(X; Y)$, удовлетворяющую равенству

$$(M; N)\cdot(X; Y)=(P; Q).$$

Из определения произведения пар и определения равных пар следует, что

$$(X; Y)=(P\cdot N; Q\cdot M).$$

О паре $(P\cdot N; Q\cdot M)$ говорят, что она получена в результате *деления* пары $(P; Q)$ на пару $(M; N)$, и пишут

$$(P; Q):(M; N)=(P\cdot N; Q\cdot M).$$

Из определения частного пар следует равенство

$$(P; Q):(M; N)=(P; Q)\cdot(N; M),$$

т. е. деление на пару $(M; N)$ может быть заменено умножением на пару $(N; M)$. Пара $(N; M)$ называется обратной паре $(M; N)$.

Пусть многочлены P и Q в упорядоченной паре $(P; Q)$ таковы, что многочлен P представим в виде

$$P=Q\cdot R,$$

где R — некоторый многочлен, т. е. многочлен P нацело делится на многочлен Q. Будем отождествлять пару таких многочленов $(P; Q)$ с многочленом R.

Пусть $(P; Q)$ и $(M; N)$ — такие две пары многочленов, что

$$P=Q\cdot R \quad \text{и} \quad M=N\cdot S, \tag{2}$$

где R и S — многочлены.

При принятом отождествлении пар $(P; Q)$ и $(M; N)$ с многочленами R и S соответственно, сумма, разность и произведение пар $(P; Q)$ и $(M; N)$ может быть отождествлена с суммой, разностью и произведением многочленов R и S.

Действительно, сумма пар $(P; Q)$ и $(M; N)$ по правилу сложения пар равна паре

$$(P\cdot N+Q\cdot M; Q\cdot N).$$

Так как $P=Q\cdot R$, а $M=N\cdot S$, то полученная сумма пар может быть записана в виде

$$(Q\cdot R\cdot N+Q\cdot N\cdot S; Q\cdot N). \tag{3}$$

Поскольку первый многочлен этой пары нацело делится на второй многочлен пары и их частным является многочлен

$$(Q\cdot R\cdot N+Q\cdot N\cdot S):(Q\cdot N)=R+S, \tag{4}$$

то пара (3) отождествляется с многочленом $R+S$.

Таким образом, если многочлены, образующие пары $(P; Q)$ и $(M; N)$, связаны равенствами (2), то сумма многочленов R и S, вычисляемая как сумма пар $(P; Q)$ и $(M; N)$, представляет собой пару, которая отождествляется с суммой многочленов $R+S$.

Аналогично проверяется, что произведение и разность многочленов R и S, рассматриваемых как пары $(P; Q)$ и $(M; N)$, при выполнении условия (2) по правилу умножения и вычитания пар определяют пары, которые можно отождествить с многочленами $N\cdot S$ и $R-S$ соответственно.

Из определения умножения и деления пар следует, что частное двух многочленов R и S, рассматриваемых как упорядоченные пары $(P; Q)$ и $(M; N)$, при условии (2) в множестве пар всегда существует.

Действительно, если многочлены P и Q записать в виде $(P; 1)$ и $(Q; 1)$ соответственно, то из равенства

$$P:Q=(P; 1):(Q; 1)=(P; 1)\cdot(1; Q)=(P; Q)$$

следует, что частное двух многочленов P и Q всегда представляется в виде пары $(P; Q)$.

В тех случаях, когда многочлен P не делится нацело на многочлен Q, пара $(P; Q)$ представляет собой не совпадающее ни с каким многочленом частное от деления многочлена P на многочлен Q, которое называется *рациональной алгебраической дробью* и обычно обозначается символом P/Q.

3.2. Правильные алгебраические дроби. Рациональная алгебраическая дробь $\frac{P(x)}{Q(x)}$ называется *правильной*, если степень многочлена в числителе дроби меньше степени многочлена в знаменателе, и *неправильной* в противном случае.

Так, дробь $\frac{x^2}{x^3+1}$ правильная, а дроби $\frac{x^3}{x^2-3}$ и $\frac{x^3}{x+1}$ неправильные.

Всякую неправильную рациональную алгебраическую дробь $\frac{P(x)}{Q(x)}$ можно представить в виде суммы многочлена и правильной дроби. Для этого следует найти частное $G(x)$ и остаток $R(x)$ от деления многочлена $P(x)$ на многочлен $Q(x)$ и записать дробь в виде

$$\frac{P(x)}{Q(x)} = G(x) + \frac{R(x)}{Q(x)}, \tag{5}$$

где $\frac{R(x)}{Q(x)}$ — правильная дробь, а $G(x)$ — многочлен, называемой *целой частью* рациональной алгебраической дроби.

Представление неправильной дроби $\frac{P(x)}{Q(x)}$ в виде (5) называется *выделением целой части* неправильной алгебраической дроби.

Пример. Рациональная алгебраическая дробь $\frac{x^3}{x+1}$ может быть записана в виде суммы многочлена и правильной рациональной алгебраической дроби:

$$\frac{x^3}{x+1} = x^2 - x + 1 - \frac{1}{x+1},$$

где x^2-x+1 — целая часть неправильной алгебраической дроби.

3.3. Простейшие дроби. Правильная рациональная алгебраическая дробь $\frac{P(x)}{Q(x)}$ называется *простейшей*, если ее знаменатель $Q(x)$ является натуральной степенью некоторого неприводимого многочлена $q(x)$ *):

$$Q(x) = q^k(x), \quad k \geqslant 1,$$

а степень числителя $P(x)$ меньше степени многочлена $q(x)$. Напомним, что среди многочленов с действительными коэффициентами и со старшим коэффициентом, равным единице, неприводимыми являются лишь линейные многочлены $x-c$ и квадратные многочлены $x^2 + px + q$ при условии, что коэффициенты квадратного трехчлена удовлетворяют неравенству $p^2 - 4q < 0$. Вследствие этого рациональная алгебраическая дробь может быть простейшей лишь в случаях, когда ее числитель $P(x)$ — либо многочлен первой степени, либо многочлен нулевой степени (т. е. число, не равное нулю).

*) О неприводимых многочленах см. п. 1.10.

Примеры. 1. Дробь $\frac{x-1}{(x^2+1)^k}$ (k натуральное) — простейшая рациональная алгебраическая дробь: ее знаменатель есть степень неприводимого многочлена, а степень неприводимого многочлена больше степени числителя.

2. Дробь $\frac{x^3-3}{x^4+1}$ не является простейшей, так как степень многочлена, стоящего в числителе дроби, больше степени многочлена, стоящего в знаменателе.

3. Дробь $\frac{x}{x^2-1}$ не является простейшей, так как ее знаменатель можно разложить на множители, т. е. знаменатель не является неприводимым многочленом.

В теории рациональных алгебраических дробей центральное место занимает следующая теорема.

Всякая правильная рациональная дробь разлагается в сумму простейших дробей, и это разложение единственно.

Пусть $\frac{P(x)}{Q(x)}$ — правильная дробь, знаменатель которой может быть представлен в виде следующего произведения натуральных степеней неприводимых многочленов:

$$Q(x) = q_1^{k_1}(x)\, q_2^{k_2}(x) \ldots q_l^{k_l}(x),$$

где $q_i(x) \neq q_j(x)$ при $i \neq j$ и $k_1, k_2, \ldots, k_l$ — натуральные числа.

Тогда

$$\frac{P(x)}{Q(x)} = \frac{p_1(x)}{q_1^{k_1}(x)} + \frac{p_2(x)}{q_2^{k_2}(x)} + \ldots + \frac{p_l(x)}{q_l^{k_l}(x)}, \tag{6}$$

где все слагаемые в правой части — правильные дроби, каждая из которых может быть представлена в виде суммы простейших дробей:

$$\frac{p(x)}{q^k(x)} = \frac{S_k(x)}{q^k(x)} + \frac{S_{k-1}(x)}{q^{k-1}(x)} + \ldots + \frac{S_2(x)}{q^2(x)} + \frac{S_1(x)}{q(x)}, \tag{7}$$

причем степени всех числителей $S_k(x)$, $S_{k-1}(x)$, ..., $S_1(x)$ в правой части разложения (7) меньше степени многочлена $q(x)$.

Пример. Разложить в сумму простейших дробей правильную дробь $\frac{P(x)}{Q(x)}$, где

$$P(x) = 2x^4 - 10x^3 + 7x^2 + 4x + 3,$$
$$Q(x) = x^5 - 2x^3 + 2x^2 - 3x + 2.$$

Многочлен $Q(x)$ может быть представлен в виде произведения неприводимых многочленов:

$$x^5 - 2x^3 + 2x^2 - 3x + 2 = (x+2)(x-1)^2(x^2+1).$$

В согласии с равенствами (6), (7) искомое разложение имеет вид

$$\frac{P(x)}{Q(x)} = \frac{A}{x+2} + \frac{B}{(x-1)^2} + \frac{C}{x-1} + \frac{Dx+E}{x^2+1},$$

где числа A, B, C, D, E еще предстоит найти. Приводя правую часть дроби к общему знаменателю, из условия равенства дробей получаем равенство

$$2x^4 - 10x^3 + 7x^2 + 4x + 3 = A(x-1)^2(x^2+1) + B(x+2)(x^2+1) + C(x+2)(x-1)(x^2+1) + Dx(x+2)(x-1)^2 + E(x+2)(x-1)^2.$$

Приравнивая коэффициенты при одинаковых степенях переменных в левой и правой частях последнего равенства, получим систему пяти линейных

уравнений с пятью неизвестными, которая имеет единственное решение:

$$A=3, \quad B=1, \quad C=-2, \quad D=1, \quad E=-3.$$

Следовательно, искомое разложение имеет вид

$$\frac{P(x)}{Q(x)}=\frac{3}{x+2}+\frac{1}{(x-1)^2}-\frac{2}{x-1}+\frac{x-3}{x^2+1}.$$

§ 4. Иррациональные алгебраические выражения

Алгебраическое выражение называют *иррациональным*, если над переменными, входящими в алгебраическое выражение, кроме операций сложения, вычитания, умножения и деления производится операция возведения в рациональную (не целую) степень.

Примеры. 1. Алгебраическое выражение

$$x^2+a\sqrt{x+b}$$

относительно переменной x является иррациональным алгебраическим выражением, так как переменная x находится под знаком квадратного корня.

2. Алгебраическое выражение

$$\sqrt{2}\,x^3-\sqrt[3]{a}\,x-1$$

относительно переменной x не будет иррациональным алгебраическим выражением. Если же переменной в данном алгебраическом выражении считать a, то оно — иррациональное. Таким образом, ответ на вопрос об иррациональности алгебраического выражения зависит от того, какие величины в данном алгебраическом выражении считаются переменными, а какие — коэффициентами.

Преобразования иррациональных алгебраических выражений выполняются в соответствии с общими законами арифметических действий и правилами действий над радикалами.

Исключение иррациональности из числителя (знаменателя) иррациональной алгебраической дроби. Алгебраическое выражение вида

$$\frac{f(x)}{g(x)} \tag{1}$$

называют *дробным иррациональным* выражением, если хотя бы одно из алгебраических выражений $f(x)$ или $g(x)$ является иррациональным относительно переменной x.

Пусть $S(x)$ — некоторое иррациональное алгебраическое выражение относительно переменной x. Алгебраическое выражение $\overline{S}(x)$, не равное тождественно нулю, называют *дополнительным множителем* для алгебраического выражения $S(x)$, если произведение $S(x)\overline{S}(x)$ — рациональное алгебраическое выражение относительно переменной x. Знание дополнительных множителей для алгебраических выражений, стоящих в числителе (или знаменателе) дроби (1), позволяет представить ее в виде дроби, у которой числитель (знаменатель) — рациональное алгебраическое выражение:

$$\frac{f}{g}=\frac{f\cdot\overline{f}}{g\cdot\overline{f}}=\frac{f\cdot\overline{g}}{g\cdot\overline{g}},$$

где $\overline{f}$ — дополнительный множитель числителя, $\overline{g}$ — дополнительный

множитель знаменателя. Такое преобразование иррационального выражения называется *исключением иррациональности* соответственно из числителя или знаменателя иррационального дробного выражения.

Таким образом, процедура исключения иррациональности из числителя или знаменателя дробного иррационального выражения сводится к нахождению дополнительного множителя для алгебраического выражения S, содержащего радикалы. В общем случае трудно указать универсальный способ нахождения дополнительного множителя для произвольного иррационального выражения. Ниже приведены дополнительные множители для некоторых простейших иррациональных выражений относительно двух переменных, полученные с помощью формул сокращенного умножения:

$S(x, y)$	$\overline{S}(x, y)$	$S(x, y)\,\overline{S}(x, y)$
$\sqrt[n]{x^k y^l}$	$\sqrt[n]{x^{n-k} y^{n-l}}$	xy
$\sqrt{x} \pm \sqrt{y}$	$\sqrt{x} \mp \sqrt{y}$	$x-y$
$\sqrt[3]{x} \pm \sqrt[3]{y}$	$\sqrt[3]{x^2} \mp \sqrt[3]{xy} + \sqrt[3]{y^2}$	$x \pm y$
$\sqrt[n]{x} - \sqrt[n]{y}$ (n – любое)	$\sqrt[n]{x^{n-1}} + \sqrt[n]{x^{n-2}y} + \ldots + \sqrt[n]{y^{n-1}}$	$x-y$
$\sqrt[n]{x} + \sqrt[n]{y}$ (n – нечетное).	$\sqrt[n]{x^{n-1}} - \sqrt[n]{x^{n-2}y} + \ldots + \sqrt[n]{y^{n-1}}$	$x+y$

Если алгебраическое выражение $\overline{S}(x, y)$ — дополнительный множитель для алгебраического выражения $S(x, y)$, то и алгебраическое выражение $S(x, y)$ будет дополнительным множителем для алгебраического выражения $\overline{S}(x, y)$.

Если число слагаемых в выражении, содержащем радикалы, больше двух, в некоторых случаях удобно исключать радикалы из числителя (знаменателя) дроби, используя формулы сокращенного умножения иррациональностей.

Примеры. 1. Исключить иррациональность из знаменателя алгебраической дроби

$$\frac{x}{\sqrt[3]{x^2} + \sqrt[3]{x} + 1}.$$

Дополнительным множителем для знаменателя служит иррациональное алгебраическое выражение $\sqrt[3]{x} - 1$. Умножая числитель и знаменатель дроби на $\sqrt[3]{x} - 1$, получаем алгебраическую дробь, знаменатель которой является рациональным алгебраическим

выражением:

$$\frac{x}{\sqrt[3]{x^2}+\sqrt[3]{x}+1}=\frac{x\left(\sqrt[3]{x}-1\right)}{\left(\sqrt[3]{x}-1\right)\left(\sqrt[3]{x^2}+\sqrt[3]{x}+1\right)}=\frac{x\left(\sqrt[3]{x}-1\right)}{x-1}.$$

2. Исключить иррациональность из знаменателя дроби

$$\frac{1}{\sqrt{x}+\sqrt{y}-\sqrt{z}}.$$

Обозначив $\sqrt{x}+\sqrt{y}=t$ и умножив числитель и знаменатель данной дроби на выражение $t+\sqrt{z}$, получим

$$\frac{1}{t-\sqrt{z}}=\frac{t+\sqrt{z}}{t^2-z}=\frac{\sqrt{x}+\sqrt{y}+\sqrt{z}}{(\sqrt{x}+\sqrt{y})^2-z}=\frac{\sqrt{x}+\sqrt{y}+\sqrt{z}}{x+y-z+2\sqrt{xy}}.$$

В результате проведенных преобразований исходной алгебраической дроби мы получили алгебраическую дробь, содержащую в знаменателе лишь один радикал. Теперь, умножая числитель и знаменатель дроби на выражение $(x+y-z)-2\sqrt{xy}$, получим дробь, уже не содержащую иррациональности в знаменателе:

$$\frac{(\sqrt{x}+\sqrt{y}-\sqrt{z})(x+y-z-2\sqrt{xy})}{(x+y-z)^2-4xy}.$$

§ 5. Уравнения. Алгебраические уравнения

5.1. Основные определения. В алгебре рассматриваются два вида равенств — тождества и уравнения. *Тождество* — это равенство, которое выполняется при всех (допустимых) значениях входящих в него букв *). Для записи тождества наряду со знаком $=$ также используется знак $\equiv$.

Уравнение — это равенство, которое выполняется лишь при некоторых значениях входящих в него букв. Буквы, входящие в уравнение, по условию задачи могут быть неравноправны: одни могут принимать все свои допустимые значения (их называют *коэффициентами*, реже — *параметрами* уравнения, и обычно обозначают первыми буквами латинского алфавита: a, b, c, ... — или теми же буквами, снабженными индексами: a_1, a_2, ... или b_1, b_2, ...); другие, значения которых требуется отыскать, называют *неизвестными* (их обычно обозначают последними буквами латинского алфавита: x, y, z, ... — или теми же буквами, снабженными индексами, например x_1, x_2, ...).

В общем виде уравнение с n неизвестными x_1, x_2, ..., x_n может быть записано как

$$F(x_1, x_2, \ldots, x_n)=0,$$

где F — некоторая функция указанных аргументов.

*) Под *допустимыми* понимаются те численные значения букв, при которых выполнимы все операции, совершаемые над буквами, входящими в равенство. Например, допустимыми значениями букв, входящих в равенство

$$\frac{b+1}{a}=\sqrt{x-2},$$

будут следующие: $a\in(-\infty;\ 0)\cup(0;\ +\infty)$; $x\in[2;\ +\infty)$, $b\in(-\infty;\ +\infty)$.

В зависимости от числа неизвестных уравнение называют *уравнением с одним, двумя* и т. д. *неизвестными*.

Областью (множеством) допустимых значений неизвестных уравнения называют область определения функции $F(x_1, x_2, \ldots, x_n)$.

Значения неизвестных, обращающие уравнение в тождество, называют *решениями* уравнения. Уравнение считается решенным, если найдены все его решения или показано, что оно решений не имеет.

Если все решения уравнения $F=0$ являются решениями уравнения $G=0$, то говорят, что уравнение $G=0$ есть *следствие* уравнения $F=0$, и пишут

$$F=0 \Rightarrow G=0.$$

Два уравнения

$$F=0 \quad \text{и} \quad G=0$$

называют *эквивалентными*, если каждое из них является следствием другого, и пишут

$$F=0 \Leftrightarrow G=0.$$

Таким образом, два уравнения считаются эквивалентными, если множества решений этих уравнений совпадают.

Уравнение $F=0$ считают эквивалентным двум (или нескольким) уравнениям $F_1=0$, $F_2=0$, если множество решений уравнения $F=0$ совпадает с объединением множеств решений уравнений $F_1=0$, $F_2=0$.

Некоторые эквивалентные уравнения.

1. Уравнение $F+G=G$ эквивалентно уравнению $F=0$, рассматриваемому на множестве допустимых значений исходного уравнения.

2. Уравнение $\frac{F}{G}=0$ эквивалентно уравнению $F=0$, рассматриваемому на множестве допустимых значений исходного уравнения.

3. Уравнение $F \cdot G=0$ эквивалентно двум уравнениям: $F=0$ и $G=0$.

4. Уравнение $F^n=0$ эквивалентно уравнению $F=0$.

5. Уравнение $F^n=G^n$ при нечетном n эквивалентно уравнению $F=G$, а при четном n эквивалентно двум уравнениям: $F=G$ и $F=-G$.

Алгебраическим уравнением называется уравнение вида

$$P_n=0,$$

где P_n — многочлен n-й степени от одной или нескольких переменных.

Алгебраическим уравнением с одним неизвестным называется уравнение, сводящееся к уравнению вида

$$a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n=0,$$

где n — неотрицательное целое число; коэффициенты многочлена a_0, a_1, a_2, $\ldots$, a_{n-1}, a_n называются *коэффициентами* (или *параметрами*) уравнения и считаются заданным; x называется *неизвестным* и является искомым. Число n называется *степенью* уравнения.

Значения неизвестного x, обращающие алгебраическое уравнение в тождество, называются *корнями* (или *решениями*) алгебраического уравнения.

Прежде чем перейти к частным видам алгебраических уравнений с действительными коэффициентами (линейным, квадратным и другим уравнениям), заметим, что методы решений линейных и квадратных уравнений были известны очень давно, в то время как методы решения уравнений третьей и четвертой степеней

были найдены лишь в XVI веке. После этого почти три столетия продолжались безуспешные попытки сделать следующий шаг, т. е. найти формулы, выражающие при помощи радикалов корни любого уравнения пятой степени через его коэффициенты. Эти попытки прекратились лишь после того, как Н. Х. Абель в двадцатых годах прошлого века доказал, что такие формулы для уравнений n-й степени при $n \geqslant 5$ заведомо не могут быть найдены.

Этот результат Абеля не исключал, однако, возможности того, что корни некоторых конкретных многочленов с числовыми коэффициентами все же каким-либо способом выражаются через коэффициенты при помощи некоторой комбинации радикалов. Полностью вопрос об условиях, при которых данное уравнение разрешимо в радикалах, был исследован Э. Галуа в 30-х годах прошлого века. В частности, им было показано, что для всякого $n \geqslant 5$ можно указать неразрешимые в радикалах уравнения n-й степени, даже с целочисленными коэффициентами. Таким будет, например, уравнение $x^5 - 4x - 2 = 0$.

5.2. Линейное уравнение. *Линейным уравнением* называется уравнение первой степени

$$ax + b = 0, \tag{1}$$

где a и b — некоторые действительные числа.

Линейное уравнение всегда имеет единственный корень $x = -\frac{b}{a}$, который находится следующим образом. Прибавляя к обеим частям уравнения (1) число $-b$, получаем уравнение

$$ax = -b,$$

эквивалентное уравнению (1). Разделив обе части последнего уравнения на число $a \neq 0$, получаем корень уравнения (1):

$$x = -\frac{b}{a}.$$

5.3. Квадратное уравнение. Алгебраическое уравнение второй степени

$$ax^2 + bx + c = 0, \tag{2}$$

где a, b, c — некоторые действительные числа, называется *квадратным уравнением*. Если $a = 1$, то квадратное уравнение (2) называется *приведенным*.

Корни квадратного уравнения вычисляются по формуле

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \tag{3}$$

где выражение $b^2 - 4ac$ называется *дискриминантом* квадратного уравнения.

При этом:

если $b^2 - 4ac > 0$, то уравнение имеет два различных действительных корня;

если $b^2 - 4ac = 0$, то уравнение имеет один действительный корень кратности 2;

если $b^2 - 4ac < 0$, то уравнение действительных корней не имеет, а имеет два комплексно сопряженных корня:

$$x_1 = -\frac{b}{2a} + \frac{\sqrt{4ac - b^2}}{2a} i, \quad x_2 = -\frac{b}{2a} - \frac{\sqrt{4ac - b^2}}{2a} i.$$

Формула (3) может быть получена в результате следующих преобразований исходного уравнения. Запишем квадратный трехчлен ax^2+bx+c, стоящий в левой части уравнения (2), в виде полного квадрата и проведем преобразования, в результате которых каждый раз будет получаться уравнение, эквивалентное исходному *):

$$a\left(x+\frac{b}{2a}\right)^2-\frac{b^2-4ac}{4a}=0 \Leftrightarrow a\left(x+\frac{b}{2a}\right)^2=\frac{b^2-4ac}{4a} \Leftrightarrow \left(x+\frac{b}{2a}\right)^2=$$
$$=\frac{b^2-4ac}{4a^2} \Leftrightarrow \sqrt{\left(x+\frac{b}{2a}\right)^2}=\sqrt{\frac{b^2-4ac}{4a^2}} \Leftrightarrow \left|x+\frac{b}{2a}\right|=\frac{\sqrt{b^2-4ac}}{2|a|}.$$

Последнее уравнение эквивалентно двум уравнениям:

$$x+\frac{b}{2a}=\frac{\sqrt{b^2-4ac}}{2|a|}. \tag{4}$$

$$x+\frac{b}{2a}=-\frac{\sqrt{b^2-4ac}}{2|a|}. \tag{4'}$$

Используя определение абсолютной величины числа, легко убедиться, что уравнения (4) и (4') эквивалентны уравнениям

$$x=-\frac{b}{2a}+\frac{\sqrt{b^2-4ac}}{2a}, \quad x=-\frac{b}{2a}-\frac{\sqrt{b^2-4ac}}{2a}.$$

Объединяя эти формулы, получаем формулу для вычисления корней квадратного уравнения:

$$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}.$$

Частными видами квадратного уравнения (2) являются:

1. Приведенное квадратное уравнение (если $a=1$)

$$x^2+px+q=0.$$

Его корни вычисляются по формуле

$$x_{1,2}=-\frac{p}{2}\pm\sqrt{\left(\frac{p}{2}\right)^2-q}. \tag{5}$$

2. Квадратное уравнение с четным вторым коэффициентом:

$$ax^2+2kx+c=0 \qquad (k\text{—целое число}).$$

Корни этого квадратного уравнения удобно вычислять по формуле

$$x_{1,2}=\frac{-k\pm\sqrt{k^2-ac}}{a}. \tag{6}$$

Формулы (5) и (6)' являются частными видами формулы (3) вычисления корней квадратного уравнения.

Корни приведенного квадратного уравнения

$$x^2+px+q=0$$

---

*) Преобразование квадратного трехчлена ax^2+bx+c к виду $a\left(x+\frac{b}{2a}\right)^2-\frac{b^2-4ac}{4a}$ называется *выделением полного квадрата*.

связаны с его коэффициентами формулами Виета

$$x_1 + x_2 = -p, \quad x_1 x_2 = q.$$

В случае, если приведенное квадратное уравнение имеет действительные корни, формулы Виета позволяют судить как о знаках, так и об относительной величине корней квадратного уравнения, а именно:

если $q > 0$, $p > 0$, то оба корня отрицательны;

если $q > 0$, $p < 0$, то оба корня положительны;

если $q < 0$, $p > 0$, то уравнение имеет корни разных знаков, причем отрицательный корень по абсолютной величине больше положительного;

если $q < 0$, $p < 0$, то уравнение имеет корни разных знаков, причем отрицательный корень по абсолютной величине меньше положительного корня.

5.4. Кубическое уравнение. *Кубическим* уравнением с действительными коэффициентами называется уравнение третьей степени

$$ax^3 + bx^2 + cx + d = 0, \tag{7}$$

где a, b, c, d — некоторые действительные числа.

Кубическое уравнение (7) заменой $x = y - \frac{b}{3a}$ приводится к «неполному» кубическому уравнению относительно y

$$y^3 + py + q = 0,$$

где

$$p = -\frac{1}{3} \cdot \left(\frac{b}{a}\right)^2 + \frac{c}{a}, \quad q = 2 \cdot \left(\frac{b}{3a}\right)^3 - \frac{bc}{3a^2} + \frac{d}{a}.$$

Корни y_1, y_2, y_3 «неполного» кубического уравнения вычисляются по формулам Кардано

$$y_1 = A + B,$$

$$y_{2,3} = -\frac{A+B}{2} \pm i \cdot \frac{A-B}{2} \cdot \sqrt{3},$$

где

$$A = \sqrt[3]{-\frac{q}{2} + \sqrt{Q}}, \quad B = \sqrt[3]{-\frac{q}{2} - \sqrt{Q}}, \quad Q = \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2.$$

Здесь в качестве A и B берутся любые значения кубических корней из соответствующих комплексных чисел, удовлетворяющие равенству $A \cdot B = -p/3$.

Если коэффициенты уравнения (7) действительны, то оно имеет или один действительный корень и два комплексно сопряженных корня, или три действительных корня, из которых по крайней мере два равны, или три различных действительных корня, в зависимости от того, будет Q соответственно положительно, равно нулю или отрицательно.

5.5. Уравнения четвертой степени. Уравнение четвертой степени

$$ax^4 + bx^3 + cx^2 + dx + k = 0$$

является последним из алгебраических уравнений (после линейных, квадратных и кубических), корни которого могут быть выражены в виде комбинации радикалов от коэффициентов a, b, c, d, k.

Уравнение четвертой степени

$$ax^4+bx^3+cx^2+dx+k=0$$

с действительными коэффициентами заменой

$$y=x+\frac{b}{4a}$$

сводится к приведенному уравнению

$$y^4+py^2+qy+r=0,$$

корни которого вычисляются по формулам

$$y_1=\frac{1}{2}(\sqrt{z_1}+\sqrt{z_2}+\sqrt{z_3}), \qquad y_2=\frac{1}{2}(\sqrt{z_1}-\sqrt{z_2}-\sqrt{z_3}),$$

$$y_3=\frac{1}{2}(-\sqrt{z_1}+\sqrt{z_2}-\sqrt{z_3}), \qquad y_4=\frac{1}{2}(-\sqrt{z_1}-\sqrt{z_2}+\sqrt{z_3}),$$

где z_1, z_2, z_3 — корни кубического уравнения (*кубической резольвенты*)

$$z^3+2pz^2+(p^2-4r)\,z-q^2=0$$

и знаки перед корнями выбираются так, чтобы выполнялось условие $\sqrt{z_1}\sqrt{z_2}\sqrt{z_3}=-q$.

5.6. Некоторые уравнения четвертой степени специального вида. Биквадратное уравнение. Алгебраическое уравнение четвертой степени

$$ax^4+bx^2+c=0,$$

где a, b, c — некоторые действительные числа, называется *биквадратным уравнением*. Заменой $x^2=y$ решение биквадратного уравнения сводится к решению квадратного уравнения $ay^2+by+c=0$ с последующим решением двух двучленных уравнений $x^2=y_1$ и $x^2=y_2$ (y_1 и y_2 — корни соответствующего квадратного уравнения).

Если $y_1\geqslant 0$ и $y_2\geqslant 0$, то биквадратное уравнение имеет четыре действительных корня

$$x_{1,2}=\pm\sqrt{y_1}, \qquad x_{3,4}=\pm\sqrt{y_2}.$$

Если $y_1\geqslant 0$, $y_1<0$ *), то биквадратное уравнение имеет два действительных корня $x_{1,2}=\pm\sqrt{y_1}$ и два чисто мнимых сопряженных корня

$$x_{3,4}=\pm i\sqrt{-y_2}.$$

Если $y_1<0$ и $y_2<0$, то биквадратное уравнение имеет четыре чисто мнимых попарно сопряженных корня

$$x_{1,2}=\pm i\sqrt{-y_1}, \qquad x_{3,4}=\pm i\sqrt{-y_2}.$$

Если y_1 и y_2 — комплексные корни соответствующего квадратного уравнения, то для нахождения четырех корней биквадратного уравнения надо извлечь квадратные корни из комплексных чисел y_1 и y_2. Этой довольно громоздкой операции можно избежать следующим образом. Запишем исходное уравнение $ax^4+bx^2+c=0$ в виде

$$x^4+px^2+q=0, \tag{8}$$

где $p=b/a$, $q=c/a$, причем коэффициенты p и q удовлетворяют условиям $|p|<2\sqrt{q}$ и $q>0$.

*) Случай $y_1<0$, $y_2\geqslant 0$ аналогичен разобранному.

Преобразуем левую часть уравнения (8):
$x^4+px^2+q=(x^4+q)+px^2=$

$$=(x^4+2\sqrt{q}x^2+q)-(2\sqrt{q}-p)x^2=(x^2+\sqrt{q})^2-(2\sqrt{q}-p)x^2=$$
$$=\left(x^2+x\sqrt{2\sqrt{q}-p}+\sqrt{q}\right)\left(x^2-x\sqrt{2\sqrt{q}-p}+\sqrt{q}\right).$$

Теперь решение биквадратного уравнения $ax^4+bx^2+c=0$ сводится к решению двух квадратных уравнений с действительными коэффициентами:

$$x^2+x\sqrt{2\sqrt{q}-p}+\sqrt{q}=0,$$
$$x^2-x\sqrt{2\sqrt{q}-p}+\sqrt{q}=0.$$

Решение уравнений вида $ax^{2k}+bx^k+c=0$ ($a\neq 0$, k — натуральное число) заменой $x^k=y$ сводится к решению квадратного уравнения $ay^2+by+c=0$ с последующим решением соответствующих двучленных уравнений.

Возвратное уравнение. Алгебраическое уравнение четвертой степени

$$ax^4+bx^3+cx^2+dx+e=0 \qquad (9)$$

при $e\neq 0$ называется *возвратным*, если коэффициенты уравнения a, b, d, e связаны равенствами $d=\lambda b$, $e=\lambda^2 a$ (λ — некоторое отличное от нуля число). Используя эту связь между коэффициентами, уравнение (9) можно записать в виде

$$ax^4+bx^3+cx^2+\lambda bx+\lambda^2 a=0. \qquad (10)$$

Поскольку $x=0$ не является корнем уравнения (9), то, разделив почленно обе части уравнения (10) на x^2 и проведя соответствующую группировку членов левой части уравнения, получим уравнение, эквивалентное уравнению (10):

$$a\left(x^2+\frac{\lambda^2}{x^2}\right)+b\left(x+\frac{\lambda}{x}\right)+c=0.$$

Теперь заменой $x+\frac{\lambda}{x}=y$ $\left(\text{учитывая, что } x^2+\frac{\lambda^2}{x^2}=y^2-2\lambda\right)$ последнее уравнение сводится к квадратному уравнению относительно y:

$$ay^2+by+c-2\lambda a=0. \qquad (11)$$

Решая уравнение (11), приходим к тому, что решение возвратного уравнения (10) сводится к решению двух квадратных уравнений:

$$x^2-y_1x+\lambda=0,$$
$$x^2-y_2x+\lambda=0,$$

где y_1 и y_2 — корни уравнения (11).

Частным случаем возвратного уравнения является *симметрическое* уравнение (соответствующее $\lambda=1$)

$$ax^4+bx^3+cx^2+bx+a=0$$

и *кососимметрическое* уравнение (соответствующее $\lambda=-1$)

$$ax^4+bx^3+cx^2-bx+a=0.$$

Заменой $x+\frac{1}{x}=y$ (для симметрического) и $x-\frac{1}{x}=y$ (для кососимметрического) эти уравнения сводятся к квадратным уравнениям относительно неизвестного y.

Уравнение четвертой степени

$$(x^2+bx+c)(x^2+bx+d)=k, \tag{12}$$

где b, c, d, k — некоторые действительные числа, заменой $x^2+bx=y$ сводится к квадратному уравнению относительно неизвестного y:

$$y^2+(c+d)y+cd-k=0. \tag{13}$$

Если уравнение (13) имеет действительные корни y_1 и y_2, то корни уравнения (12) отыскиваются как корни двух квадратных уравнений с действительными коэффициентами:

$$x^2+bx-y_1=0, \quad x^2+bx-y_2=0.$$

Решение уравнения

$$x(x+a)(x+b)(x+a+b)=c, \tag{14}$$

где a, b, c — некоторые действительные числа, может быть сведено к решению двух квадратных уравнений следующим образом. Перемножая первый и четвертый, второй и третий сомножители, получаем уравнение

$$[x^2+(a+b)x][x^2+(a+b)x+ab]=c,$$

которое заменой $x^2+(a+b)x=y$ сводится к квадратному уравнению относительно нового неизвестного y:

$$y^2+aby-c=0. \tag{15}$$

Если уравнение (15) имеет действительные корни y_1 и y_2, то множество корней уравнения (14) находится как множество корней следующих двух квадратных уравнений с действительными коэффициентами:

$$x^2+(a+b)x-y_1=0, \quad x^2+(a+b)x-y_2=0.$$

5.7. Двучленные уравнения. Уравнение n-й степени

$$ax^n \pm b=0 \tag{16}$$

называется *двучленным уравнением*. При $a>0$ и $b>0$ заменой *) $x=y\sqrt[n]{\frac{b}{a}}$, где $\sqrt[n]{\frac{b}{a}}$ — арифметическое значение корня, уравнение (16) приводится к уравнению

$$y^n \pm 1=0,$$

которое и будет далее рассматриваться.

Двучленное уравнение $y^n-1=0$ при нечетном n имеет один действительный корень $y=1$. В множестве комплексных чисел это уравнение имеет n корней (один действительный и $n-1$ комплексных):

$$y_k=\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n}, \quad k=0,\ 1,\ 2,\ \ldots,\ n-1. \tag{17}$$

Двучленное уравнение $y^n-1=0$ при четном n в множестве действительных чисел имеет два корня $\{y=1;\ y=-1\}$, а в множестве комплексных чисел имеет n корней, вычисляемых по формуле (17).

---

*) Если a и b имеют разные знаки, то $x=y\sqrt[n]{\left|\frac{b}{a}\right|}$.

Двучленное уравнение $y^n+1=0$ при нечетном n имеет один действительный корень $y=-1$, а в множестве комплексных чисел имеет n корней:

$$y_k=\cos\frac{2\pi k+\pi}{n}+i\sin\frac{2\pi k+\pi}{n},\quad k=0,\ 1,\ 2,\ \ldots,\ n-1. \qquad (18)$$

Двучленное уравнение $y^n+1=0$ при четном n действительных корней не имеет. В множестве комплексных чисел уравнение имеет n корней, вычисляемых по формуле (18).

Корни двучленного уравнения для некоторых конкретных значений n:

1) $y^2-1=0$ ($n=2$). Уравнение имеет два действительных корня $y_{1,\,2}=\pm 1$.

2) $y^3-1=0$ ($n=3$). Уравнение имеет один действительный корень $y_1=1$ и два комплексных корня $y_{2,\,3}=\dfrac{-1\pm i\sqrt{3}}{2}$.

3) $y^4-1=0$ ($n=4$). Уравнение имеет два действительных корня $y_{1,\,2}=\pm 1$ и два комплексных корня $y_{3,\,4}=\pm i$.

4) $y^2+1=0$ ($n=2$). Уравнение действительных корней не имеет. Комплексные корни: $y_{1,\,2}=\pm i$.

5) $y^3+1=0$ ($n=3$). Уравнение имеет один действительный корень $y_1=-1$ и два комплексных корня $y_{2,\,3}=\dfrac{1\pm i\sqrt{3}}{2}$.

6) $y^4+1=0$ ($n=4$). Уравнение действительных корней не имеет. Комплексные корни:

$$y_{1,\,2}=\frac{\sqrt{2}}{2}(1\pm i),\qquad y_{3,\,4}=-\frac{\sqrt{2}}{2}(1\pm i).$$

5.8. Решение алгебраического уравнения с целыми коэффициентами. Рациональные корни алгебраического уравнения n-й степени

$$a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n=0, \qquad (19)$$

где a_0, a_1, a_2, ..., a_{n-1}, a_n — целые числа, можно найти, используя следующую теорему.

Рациональными корнями уравнения (19) *могут быть лишь числа* m/p (m *целое*, p *натуральное*), *где число* $|m|$ *является делителем числа* $|a_n|$, *а число* p *— делителем числа* $|a_0|$.

Пример. Найти корни уравнения

$$4x^4+8x^3-3x^2-7x+3=0. \qquad (20)$$

Делителями числа 3 будут числа 1, 3, а делителями числа 4 — числа 1, 2, 4. Множеством значений m будет $\{1;\ -1;\ 3;\ -3\}$, а множеством значений p будет $\{1;\ 2;\ 4\}$. Множеством всевозможных различных рациональных чисел будет $\{\pm 1;\ \pm 3;\ \pm 1/2;\ \pm 1/4;\ \pm 3/2;\ \pm 3/4\}$. Непосредственная подстановка этих чисел в уравнение (20) показывает, что числа $1/2$ и $-3/2$ — корни данного уравнения и, следовательно, данный многочлен делится на линейные многочлены

$$\left(x-\frac{1}{2}\right)\quad\text{и}\quad\left(x+\frac{3}{2}\right),$$

а также и на их произведение

$$\left(x-\frac{1}{2}\right)\left(x+\frac{3}{2}\right)=x^2+x-\frac{3}{4}.$$

Произведя деление «уголком», находим многочлен частного:

$$4x^2+4x-4.$$

Решая квадратное уравнение

$$4x^2+4x-4=0,$$

получим еще два действительных корня уравнения (20):

$$x_3=\frac{-1+\sqrt{5}}{2}, \quad x_4=\frac{-1-\sqrt{5}}{2}.$$

Итак, задача полностью решена — найдены все четыре корня исходного уравнения:

$$x_1=\frac{1}{2}, \quad x_2=-\frac{3}{2}, \quad x_3=\frac{-1+\sqrt{5}}{2}, \quad x_4=\frac{-1-\sqrt{5}}{2}.$$

5.9. Рациональные алгебраические уравнения. *Рациональным* алгебраическим уравнением называется уравнение вида

$$\frac{P(x)}{Q(x)}=0, \tag{21}$$

где $P(x)$ и $Q(x)$ — многочлены. Далее для определенности будем полагать, что $P(x)$ — многочлен m-й степени, а $Q(x)$ — многочлен n-й степени.

Множество допустимых значений рационального алгебраического уравнения (21) определяется условием $Q(x)\neq 0$, откуда следует, что $x\neq c_1$, $x\neq c_2$, ..., $x\neq c_n$, где c_1, c_2, ..., c_n — корни многочлена $Q(x)$.

Метод решения уравнения (21) заключается в следующем. Решаем уравнение

$$P(x)=0,$$

корни которого обозначим через x_1, x_2, x_3, ..., x_m. Сравниваем множества корней многочленов $P(x)$ и $Q(x)$. Те корни многочлена $P(x)$, которые не являются корнями многочлена $Q(x)$, являются корнями (решениями) рационального уравнения (21).

Пример. Найти действительные корни уравнения

$$\frac{P(x)}{Q(x)}=0,$$

где $P(x)=x^4-1$, $Q(x)=x-1$.

Многочлен $P(x)$ имеет два действительных корня (оба простые): $x_1=1$, $x_2=-1$. Многочлен $Q(x)$ имеет один корень $c_1=1$. Следовательно, уравнение имеет один действительный корень $x=-1$.

Решая то же самое уравнение в множестве комплексных чисел, получим, что уравнение $\frac{P(x)}{Q(x)}$ имеет, кроме указанного действительного корня, два комплексно сопряженных корня:

$$x_2=i, \quad x_3=-i.$$

5.10. Иррациональные уравнения. Уравнение, содержащее неизвестное (либо рациональное алгебраическое выражение от неизвестного) под знаком радикала, называют *иррациональным уравнением*. В элементарной математике иррациональные уравнения решаются в множестве действительных чисел.

Всякое иррациональное уравнение с помощью элементарных преобразований (умножения, деления, возведения в целую степень обеих частей уравнения) может быть сведено к рациональному алгебраическому уравнению. При этом следует иметь в виду, что полученное рациональное алгебраическое уравнение может оказаться не эквивалентным исходному иррациональному уравнению, а именно может содержать «лишние» корни, которые не будут корнями исходного иррационального уравнения. Поэтому, вычислив корни полученного алгебраического уравнения, необходимо проверить, будут ли все они также и корнями исходного иррационального уравнения.

В общем случае трудно указать какой-либо универсальный метод решения любого иррационального уравнения, так как желательно, чтобы в результате преобразований исходного иррационального уравнения получилось не просто какое-то рациональное алгебраическое уравнение, среди корней которого будут и корни данного иррационального уравнения, а рациональное алгебраическое уравнение, образованное из многочленов как можно меньшей степени. Желание получить рациональное алгебраическое уравнение, образованное из многочленов как можно меньшей степени, вполне естественно, так как нахождение всех корней рационального алгебраического уравнения само по себе может оказаться довольно трудной задачей, решать которую полностью мы можем лишь в весьма ограниченном числе случаев.

Приведем некоторые стандартные методы решения иррациональных алгебраических уравнений.

1. Одним из самых простых приемов решения иррациональных уравнений является метод освобождения от радикалов путем последовательного возведения обеих частей уравнения в соответствующую натуральную степень. При этом если обе части уравнения возводятся в нечетную степень, то получается уравнение, эквивалентное исходному; если же обе части уравнения возводятся в четную степень, то полученное уравнение может оказаться не эквивалентным исходному уравнению. В этом легко убедиться, возведя обе части уравнения

$$f(x) = g(x)$$

в любую четную степень $2n$. В результате этой операции получается уравнение

$$|f(x)|^{2n} = |g(x)|^{2n},$$

множество решений которого представляет собой объединение множеств решений двух уравнений

$$f(x) = g(x) \quad \text{и} \quad f(x) = -g(x).$$

Однако, несмотря на этот недостаток, именно процедура возведения обеих частей уравнения в некоторую (часто четную) степень является самой распространенной процедурой сведения иррационального уравнения к рациональному уравнению.

Пример 1. Решить уравнение

$$\sqrt{P(x)} + \sqrt{Q(x)} = R(x), \tag{22}$$

где $P(x)$, $Q(x)$, $R(x)$ — некоторые многочлены.

В силу определения операции извлечения корня в множестве действительных чисел допустимые значения неизвестного x определяются условиями

$$P(x) \geqslant 0, \quad Q(x) \geqslant 0.$$

Возведя обе части уравнения (22) в квадрат, получим уравнение

$$2\sqrt{P(x)\,Q(x)} = R^2(x) - P(x) - Q(x).$$

Повторное возведение обеих частей уравнения в квадрат приводит к рациональному алгебраическому уравнению

$$4P(x)\,Q(x) = [R^2(x) - P(x) - Q(x)]^2. \tag{23}$$

Поскольку обе части уравнения (22) возводились в квадрат, может оказаться, что не все корни уравнения (23) будут решениями исходного уравнения и необходима проверка корней.

2. Другим приемом решения иррациональных уравнений является способ введения новых неизвестных, относительно которых получается либо более простое иррациональное уравнение, либо рациональное уравнение.

Пример 2. Решить иррациональное уравнение

$$(3-x)\sqrt[3]{\frac{3-x}{x-1}}+(x-1)\sqrt[3]{\frac{x-1}{3-x}}=2.$$

Находим множество допустимых значений для этого уравнения:

$$x\in(-\infty;\ 1)\cup(1;\ 3)\cup(3;\ +\infty).$$

Положив $\sqrt[3]{\frac{3-x}{x-1}}=y$, после подстановки получим уравнение

$$(3-x)\,y+\frac{x-1}{y}=2$$

или эквивалентное ему уравнение

$$(3-x)\,y^2-2y+x-1=0,$$

которое можно рассматривать как квадратное уравнение относительно y. Решая это уравнение, получим

$$y_1=1,\quad y_2=\frac{x-1}{3-x}.$$

Следовательно, множество решений исходного иррационального уравнения представляет собой объединение множеств решений следующих двух уравнений:

$$\sqrt[3]{\frac{3-x}{x-1}}=1,\quad \sqrt[3]{\frac{3-x}{x-1}}=\frac{x-1}{3-x}.$$

Возведя обе части каждого из этих уравнений в куб, получим два рациональных алгебраических уравнения:

$$\frac{3-x}{x-1}=1,\quad \frac{3-x}{x-1}=\left(\frac{x-1}{3-x}\right)^3.$$

Решая эти уравнения, находим, что данное иррациональное уравнение имеет единственный корень $x=2$.

В заключение заметим, что при решении иррациональных уравнений не следует начинать решение уравнения с возведения обеих частей уравнений в натуральную степень, пытаясь свести решение иррационального уравнения к решению рационального алгебраического уравнения. Сначала необходимо посмотреть, нельзя ли сделать какое-либо тождественное преобразование уравнения, которое может существенно упростить его решение.

Пример 3. Решить уравнение

$$a\sqrt[4]{1+x}+\frac{a}{x}\sqrt[4]{1+x}=\sqrt[4]{x}. \tag{24}$$

Находим множество допустимых значений: $x\in(0;\ +\infty)$. Сделаем следующие преобразования:

$$a\sqrt[4]{1+x}+\frac{a}{x}\sqrt[4]{1+x}=\sqrt[4]{x}\Leftrightarrow a\sqrt[4]{1+x}\left(1+\frac{1}{x}\right)=$$

$$=\sqrt[4]{x}\Leftrightarrow a\sqrt[4]{\frac{1+x}{x}}\left(1+\frac{1}{x}\right)=1.$$

Теперь мы можем записать уравнение в виде

$$a\left(1+\frac{1}{x}\right)^{5/4}=1,$$

тогда:

при $a=0$ уравнение решений не имеет;

при $a\neq 0$ уравнение может быть записано в виде

$$\left(1+\frac{1}{x}\right)^{5/4}=\frac{1}{a};$$

при $a<0$ уравнение решений не имеет, так как при любом $x\in(0;\ +\infty)$ левая часть уравнения положительна;

при $a>0$ получаем

$$x=\frac{1}{\sqrt[5]{1/a^4-1}}.$$

Принимая во внимание, что множество допустимых значений неизвестного определяется условием $x>0$, из условия

$$\frac{1}{\sqrt[5]{1/a^4-1}}>0$$

находим значения параметра a, при которых найденное значение неизвестного x является решением уравнения (24). Такими будут значения $a\in(0;\ 1)$. При всех остальных значениях a уравнение решений не имеет.

5.11. Уравнения, содержащие неизвестное под знаком абсолютной величины. Эти уравнения можно свести к уравнениям, не содержащим знака абсолютной величины, используя ее определение. Так, решение уравнения

$$x^2-5\left|x-\frac{5}{2}\right|-\frac{13}{2}=0 \tag{25}$$

сводится к решению двух уравнений с дополнительными условиями.

1. Если $x-\frac{5}{2}\geqslant 0$, то уравнение (25) приводится к виду

$$x^2-5x+6=0, \tag{26}$$

Решения этого уравнения: $x_1=2$, $x_2=3$. Условию $x-\frac{5}{2}\geqslant 0$ удовлетворяет лишь второй корень квадратного уравнения (26), и, следовательно, число 3 является корнем исходного уравнения (25).

2. Если $x-\frac{5}{2}<0$, то уравнение (25) приводится к виду

$$x^2+5x-19=0.$$

Корнями этого уравнения будут числа $x_1=\frac{-5+\sqrt{101}}{2}$ и $x_2=$ $=\frac{-5-\sqrt{101}}{2}$. Первый корень $x_1=\frac{-5+\sqrt{101}}{2}$ не удовлетворяет условию $x-\frac{5}{2}<0$ и поэтому не является решением уравнения (25),

Таким образом, решениями уравнения (25) будут числа 3 и $\frac{-5-\sqrt{101}}{2}$.

Заметим, что коэффициенты уравнения, содержащего неизвестное под знаком абсолютной величины, можно подобрать так, что решениями будут все значения неизвестного, принадлежащие некоторому промежутку числовой оси. Например, решим уравнение

$$|x|+|3-x|=3. \tag{27}$$

Отметим на числовой оси Ox точки 0 и 3 (нули функций, стоящих под знаком абсолютной величины). Эти точки разобьют числовую ось на три промежутка (рис. 4.1):

$$-\infty < x < 0; \quad 0 \leqslant x \leqslant 3, \quad 3 < x < +\infty.$$

На этих промежутках:

1) при $x \in (-\infty; 0)$ уравнение (27) приводится к виду $3-2x=3$ и в промежутке $(-\infty; 0)$ решений не имеет. Аналогично при

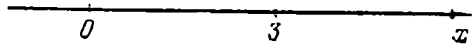


Рис. 4.1

$x \in (3; +\infty)$ уравнение (27) приводится к виду $2x-3=3$ и в промежутке $(3; +\infty)$ решений не имеет;

2) при $x \in [0; 3]$ уравнение (27) приводится к виду $x+(3-x)=3$, т. е. обращается в тождество. Следовательно, любое значение $x \in [0; 3]$ является решением уравнения (27).

5.12. Решение уравнений в множестве комплексных чисел. Рассмотрим два примера решения уравнений в множестве комплексных чисел.

Пример 1. Найти комплексные числа z, удовлетворяющие уравнению *)

$$z|z|+az+i=0, \tag{28}$$

где a — некоторое положительное действительное число.

По определению любое комплексное число z может быть записано в виде

$$z=x+iy, \tag{29}$$

где x, y — действительные числа.

По определению модуля комплексного числа

$$|z|=\sqrt{x^2+y^2}. \tag{30}$$

Подставив в уравнение (28) вместо z и $|z|$ их выражения через x и y, заданные формулами (29) и (30), получим следующую запись уравнения (28):

$$x\sqrt{x^2+y^2}+ax+i\,[y\sqrt{x^2+y^2}+ay+1]=0+0i. \tag{31}$$

Из условия равенства двух комплексных чисел следует, что уравнение (31) эквивалентно системе двух уравнений

$$x\sqrt{x^2+y^2}+ax=0, \quad y\sqrt{x^2+y^2}+ay+1=0, \tag{32}$$

решения которой отыскиваются уже в множестве действительных чисел.

Нетрудно заметить, что множество решений первого уравнения системы (32) может быть найдено как объединение множеств решений двух уравнений:

$$x=0 \text{ и } \sqrt{x^2+y^2}+a=0.$$

*) Здесь мы будем обозначать неизвестную величину не буквой x, а буквой z, подчеркивая тем самым, что решения уравнения отыскиваются в множестве комплексных чисел.

Второе уравнение решений не имеет, так как по условию задачи $a > 0$.

Подставляя $x=0$ во второе уравнение системы (32), получаем уравнение для y*):

$$y|y|+ay+1=0,$$

множество решений которого получается как объединение множеств решений двух систем

$$\begin{cases} y \geqslant 0, \\ y^2+ay+1=0; \end{cases} \quad (33) \qquad \begin{cases} y < 0, \\ -y^2+ay+1=0. \end{cases} \quad (34)$$

Нетрудно убедиться в том, что с учетом условия $a > 0$ система (33) решений не имеет, а система (34) имеет единственное решение

$$y=\frac{a-\sqrt{a^2+4}}{2}.$$

Таким образом, решением уравнения (28) является чисто мнимое число

$$z=i\,\frac{a-\sqrt{a^2+4}}{2}.$$

Пример 2. Найти комплексные числа z, удовлетворяющие уравнению

$$\bar{z}=z^{n-1}, \quad (35)$$

где $\bar{z}$ — комплексное число, сопряженное z, а n — натуральное.

Для комплексно сопряженных чисел z и $\bar{z}$ имеем $|z|=|\bar{z}|$, поэтому из уравнения (35) следует уравнение

$$|z|=|z|^{n-1}. \quad (36)$$

Это алгебраическое уравнение $(n-1)$-й степени относительно неизвестного $|z|$. Множество решений этого уравнения следующее:

1) при $n=1$ $|z|=1$;

2) при $n \neq 1$ $(n \in \mathbb{N})$ $|z|=1$ и $|z|=0$.

Рассмотрим случай $|z|=1$. Умножив обе части уравнения (35) на z, получим уравнение

$$z^n=1, \quad (37)$$

эквивалентное уравнению (35), так как $|z| \neq 0$, а значит, и $z \neq 0$.

Решениями уравнения (37) будут комплексные числа

$$z_k=\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n}, \quad k=0, 1, 2, \ldots, n-1.$$

Таким образом, решения уравнения (35) можно записать в виде:

1) $z=\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n}$ $(k=0, 1, 2, \ldots, n-1)$ при любом натуральном n;

2) $z=0$ при любом натуральном $n \neq 1$.

5.13. Диофантовы уравнения. Алгебраическое уравнение с несколькими неизвестными, коэффициенты которого — целые числа и решения которого отыскиваются в множестве целых чисел, называется *диофантовым уравнением*.

Диофантовы уравнения могут: не иметь решений; иметь конечное число решений; иметь бесконечно много решений.

Простейшее диофантово уравнение — линейное уравнение с двумя неизвестными:

$$ax+by=c, \quad (38)$$

где a, b, c — целые числа. Полное решение этого уравнения может быть найдено с помощью следующего следствия алгоритма Евклида (см. п. 1.5 гл. 2).

Если число d — наибольший общий делитель целых чисел a и b, то существуют такие целые числа k и l, что $d=ka+lb$.

*) Здесь $|y|$ — обозначение абсолютной величины действительного числа y.

Линейное диофантово уравнение (38) не имеет решений, если числа c и d взаимно простые. Если же число c кратно числу d (т. е. представимо в виде $c=pd$, где p — целое число, отличное от нуля), то одно из решений уравнения (38) имеет вид

$$x^*=pk, \quad y^*=pl.$$

Множество всех решений диофантова уравнения (38) задается формулами

$$x=x^*+\frac{nb}{d}, \quad y=y^*-\frac{na}{d}, \quad n\in Z.$$

§ 6. Трансцендентные уравнения

Уравнение, не сводящееся к алгебраическому уравнению с помощью алгебраических преобразований, называется *трансцендентным уравнением* *). Простейшими трансцендентными уравнениями являются показательные, логарифмические и тригонометрические уравнения.

6.1. Показательные уравнения. *Показательным уравнением* называется уравнение, в котором неизвестное входит только в показатели степеней при некоторых постоянных основаниях.

Простейшим показательным уравнением является уравнение

$$a^{f(x)}=b, \tag{1}$$

где a и b — некоторые положительные числа ($a\neq 1$), а $f(x)$ — некоторое алгебраическое выражение. Показательное уравнение (1) эквивалентно алгебраическому уравнению

$$f(x)=\log_a b.$$

В случае, когда $f(x)=x$, показательное уравнение (1) имеет решение

$$x=\log_a b.$$

Множество решений показательного уравнения

$$P(a^x)=0, \tag{2}$$

где P — некоторый многочлен, находится следующим образом. Вводится новое неизвестное $y=a^x$, и уравнение (2) решается как алгебраическое относительно y. После этого решение исходного уравнения (2) сводится к решению простейших показательных уравнений вида (1).

Пример 1. Решить уравнение

$$9\cdot 8^x-18\cdot 4^x-2\cdot 2^x+4=0.$$

Запишем уравнение в виде

$$9(2^x)^3-18(2^x)^2-2\cdot 2^x+4=0$$

и введем новое неизвестное $y=2^x$; получим кубическое уравнение относительно y:

$$9y^3-18y^2-2y+4=0.$$

---

*) Под *алгебраическими преобразованиями* уравнения $F=0$ понимают следующие преобразования:

1) прибавление к обеим частям уравнения одного и того же алгебраического выражения;

2) умножение обеих частей уравнения на одно и то же алгебраическое выражение;

3) возведение обеих частей уравнения в рациональную степень.

Нетрудно убедиться (см. п. 5.7 гл. 4), что это уравнение имеет единственный рациональный корень $y_1=2$ и два иррациональных корня: $y_2=\frac{\sqrt{2}}{3}$ и $y_3=-\frac{\sqrt{2}}{3}$. Таким образом, решение исходного уравнения сведено к решению простейших показательных уравнений

$$2^x=2, \quad 2^x=\frac{\sqrt{2}}{3}, \quad 2^x=-\frac{\sqrt{2}}{3}.$$

Последнее из перечисленных уравнений решений не имеет. Решениями первого и второго уравнений являются следующие значения x:

$$x=1 \quad \text{и} \quad x=\log_2\frac{\sqrt{2}}{3}.$$

Некоторые простейшие показательные уравнения:

1) уравнение

$$\alpha a^{2x}+\beta a^x+\gamma=0$$

заменой $a^x=y$ сводится к квадратному уравнению

$$\alpha y^2+\beta y+\gamma=0;$$

2) уравнение

$$\alpha a^x+\beta a^{-x}+\gamma=0$$

заменой $a^x=y$ сводится к квадратному уравнению

$$\alpha y^2+\gamma y+\beta=0;$$

3) уравнение

$$\alpha a^{2x}+\beta\,(ab)^x+\gamma b^{2x}=0$$

заменой $\left(\frac{a}{b}\right)^x=y$ сводится к квадратному уравнению

$$\alpha y^2+\beta y+\gamma=0.$$

6.2. Логарифмические уравнения. *Логарифмическим* называется уравнение, в котором неизвестным является аргумент логарифмической функции *).

Простейшим логарифмическим уравнением, решение которого сводится к решению алгебраического уравнения, является уравнение

$$\log_a f(x)=b, \tag{3}$$

где a — некоторое положительное число, отличное от единицы, b — любое действительное число, а $f(x)$ — некоторое алгебраическое выражение. Логарифмическое уравнение (3) эквивалентно алгебраическому уравнению

$$f(x)=a^b.$$

В случае, когда $f(x)=x$, логарифмическое уравнение (3) имеет решение

$$x=a^b.$$

Множество решений логарифмического уравнения вида $P(\log_a x)=0$, где P — некоторый многочлен, находится следующим образом.

*) К логарифмическим уравнениям относятся также уравнения, в которых основание логарифма зависит от неизвестной величины.

Вводится новое неизвестное $y=\log_a x$, и уравнение решается как алгебраическое уравнение относительно y. После этого решаются простейшие логарифмические уравнения вида (3).

Пример 1. Решить уравнение

$$\log_2^2 x+\log_2 x-2=0. \tag{4}$$

Относительно неизвестного $y=\log_2 x$ данное уравнение — квадратное:

$$y^2+y-2=0.$$

Корни этого уравнения: $y_1=1$, $y_2=-2$. Решая алгебраические уравнения

$$\log_2 x=1, \quad \log_2 x=-2,$$

получаем решения логарифмического уравнения (4): $x_1=2$, $x_2=1/4$.

В некоторых случаях, для того чтобы свести решение логарифмического уравнения к последовательному решению алгебраического и простейших логарифмических уравнений, необходимо предварительно сделать подходящие преобразования логарифмов, входящих в уравнение. Такими преобразованиями могут быть преобразование суммы логарифмов двух величин в логарифм произведения этих величин, переход от логарифма с одним основанием к логарифму с другим основанием и т. д.

Пример 2. Решить уравнение

$$2\sqrt[3]{2\log_{10}^2 x}-\sqrt[3]{\log_2 x}-6=0. \tag{5}$$

Для того чтобы свести решение данного уравнения к последовательному решению алгебраического и простейших логарифмических уравнений, необходимо прежде всего привести все логарифмы к одному основанию (здесь, например, к основанию 2). Для этого воспользуемся формулой

$$\log_a N=\frac{\log_b N}{\log_b a},$$

в силу которой $\log_{10} x=\frac{\log_2 x}{\log_2 10}=\frac{\log_2 x}{4}$. Подставив в уравнение (5) вместо $\log_{10} x$ равную ему величину $\frac{\log_2 x}{4}$, получаем уравнение

$$\sqrt[3]{\log_2^2 x}-\sqrt[3]{\log_2 x}-6=0.$$

Заменой $\sqrt[3]{\log_2 x}=y$ это уравнение сводится к квадратному уравнению относительно y:

$$y^2-y-6=0.$$

Корни этого квадратного уравнения: $y_1=3$, $y_2=-2$. Решаем уравнения $\sqrt[3]{\log_2 x}=3$ и $\sqrt[3]{\log_2 x}=-2$:

$$\sqrt[3]{\log_2 x}=3 \Leftrightarrow \log_2 x=27 \Leftrightarrow x=2^{27},$$

$$\sqrt[3]{\log_2 x}=-2 \Leftrightarrow \log_2 x=-8 \Leftrightarrow x=2^{-8}.$$

Пример 3. Решить уравнение

$$\log_3(2x-1)-\log_3(x-1)=1.$$

Преобразуя разность логарифмов двух величин в логарифм частного этих величин:

$$\log_3(2x-1)-\log_3(x-1)=\log_3\frac{2x-1}{x-1},$$

сводим данное уравнение к простейшему логарифмическому уравнению:

$$\log_3 \frac{2x-1}{x-1} = 1 \Leftrightarrow \frac{2x-1}{x-1} = 3 \Leftrightarrow x = 2.$$

Уравнение

$$\log_{g(x)} f(x) = c,$$

где c — число, а $f(x)$ и $g(x)$ — некоторые функции переменной x, эквивалентно уравнению

$$f(x) = [g(x)]^c,$$

рассматриваемому на множестве допустимых значений исходного уравнения, которое определяется как множество решений системы неравенств

$$f(x) > 0, \quad g(x) > 0, \quad g(x) \neq 1.$$

Пример 4. Решить уравнение

$$\log_{x-1}(2x^2 - 8x + 9) = 2. \tag{6}$$

Множество допустимых значений x определяется как множество решений системы

$$2x^2 - 8x + 9 > 0, \quad x - 1 > 0, \quad x - 1 \neq 1.$$

Первое неравенство системы выполняется при любых значениях x, и, следовательно, решениями системы будут значения $x \in (1; 2) \cup (2; +\infty)$.

Логарифмическое уравнение (6) эквивалентно уравнению $2x^2 - 8x + 9 = (x-1)^2$, корнями которого являются $x_1 = 4$, $x_2 = 2$. Множеству допустимых значений уравнения (6) принадлежит лишь первый корень ($x_1 = 4$), который и есть единственное решение данного логарифмического уравнения.

§ 7. Системы уравнений

7.1. Основные определения. *Решением* некоторого множества (системы) уравнений

$$F_1(x_1, x_2, \ldots, x_n) = 0, \ \ldots, \ F_m(x_1, x_2, \ldots, x_n) = 0$$

с неизвестными $x_1, x_2, \ldots, x_n$ называется множество значений неизвестных, обращающих одновременно все уравнения системы в тождества.

Система уравнений считается решенной, если найдены все такие значения неизвестных или доказано, что не существует набора значений неизвестных, обращающих одновременно все уравнения системы в тождества. В последнем случае говорят, что система не имеет решений или что она *несовместна*.

Систему уравнений иногда записывают, объединяя уравнения фигурной скобкой.

Две системы уравнений считают *эквивалентными*, если они имеют одно и то же множество решений. (Две несовместные системы по определению считаются эквивалентными.)

Говорят, что система уравнений (S) *эквивалентна двум системам* уравнений (S_1) и (S_2), если множество решений системы (S) совпадает с объединением множеств решений систем (S_1) и (S_2).

Свойства систем уравнений:

1) при замене любого уравнения системы эквивалентным уравнением получается эквивалентная система;

2) если одно из уравнений системы (S) эквивалентно некоторым двум уравнениям, то исходная система (S) эквивалентна двум системам (S_1) и (S_2), в каждой из которых это уравнение заменено на

одно из уравнений эквивалентной совокупности, а остальные оставлены без изменений.

Например, система

$$x^3+y^3=3(x+y),$$
$$x^2+y^2=7,$$

в которой первое уравнение эквивалентно двум уравнениям: $x+y=0$ и $x^2-xy+y^2=3$, эквивалентна двум системам:

$$x+y=0, \quad x^2-xy+y^2=3,$$
$$x^2+y^2=7, \quad x^2+y^2=7.$$

7.2. Системы линейных уравнений. *Системой s линейных алгебраических уравнений* с n неизвестными $x_1, x_2, \ldots, x_n$ называется система вида

$$\begin{aligned} &a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1,\\ &a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ &a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=b_s. \end{aligned} \tag{1}$$

Величины $a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, a_{22}, \ldots, a_{2n}, \ldots, a_{s1}, a_{s2}, \ldots, a_{sn}$ называются *коэффициентами* линейной системы уравнений. Индексы у коэффициентов линейной системы означают следующее: первый индекс указывает номер уравнения системы в записи (1), второй индекс указывает номер неизвестного, при которой стоит данный коэффициент. Так, a_{25} — коэффициент, стоящий во втором уравнении системы (1) при неизвестном x_5.

Величины $b_1, b_2, \ldots, b_s$ называются *свободными членами* первого, второго, ..., s-го уравнений системы (1). Система уравнений (1) называется *однородной*, если все числа b_i равны нулю ($i=1, 2, 3, \ldots, s$), и *неоднородной*, если хотя бы одно b_i отлично от нуля.

Упорядоченное множество n чисел $k_1, k_2, k_3, \ldots, k_n$ называется *решением системы* (1), если при подстановке его в систему вместо неизвестных $x_1, x_2, x_3, \ldots, x_n$ все уравнения системы обращаются в тождества. Система уравнений называется *совместной*, если она имеет хотя бы одно решение, и *несовместной*, если она не имеет ни одного решения.

Совместная система линейных уравнений называется *определенной*, если она имеет единственное решение, т. е. существует только один набор n чисел $k_1, k_2, \ldots, k_n$, который обращает все уравнения системы в тождества.

Совместная система линейных уравнений называется *неопределенной*, если она имеет более одного решения.

Система однородных уравнений всегда имеет нулевое решение:

$$x_1=x_2=x_3=\ldots=x_n=0.$$

Если система однородных уравнений имеет ненулевое решение $k_1, k_2, \ldots, k_n$ (т. е. хотя бы одно из чисел k_i ($i=1, 2, \ldots, n$) отлично от нуля), то она имеет бесконечно много решений вида $lk_1, lk_2, \ldots, lk_n$, где l — любое число.

7.3. Метод последовательного исключения неизвестных (метод Гаусса). Одним из наиболее распространенных методов решения систем линейных алгебраических уравнений является метод последовательного исключения неизвестных — *метод Гаусса*. Этот метод основан на преобразованиях системы линейных уравнений, в результате которых получается система, эквивалентная исходной. Проведем эти

преобразования на примере системы трех линейных уравнений с тремя неизвестными

$$\begin{aligned} a_{11}x+a_{12}y+a_{13}z&=b_1,\\ a_{21}x+a_{22}y+a_{23}z&=b_2,\\ a_{31}x+a_{32}y+a_{33}z&=b_3. \end{aligned} \tag{2}$$

1. Если обе части какого-либо уравнения системы домножить на одно и то же (не равное нулю) число, то полученная система будет эквивалентна первоначальной (т. е. они или обе несовместны, или же обе совместны и множества их решений совпадают).

Например, система

$$\begin{aligned} a_{11}x+a_{12}y+a_{13}z&=b_1,\\ ca_{21}x+ca_{22}y+ca_{23}z&=cb_2, \quad c\neq 0,\\ a_{31}x+a_{32}y+a_{33}z&=b_3 \end{aligned}$$

эквивалентна системе (2).

2. Если обе части какого-либо уравнения системы, умноженные на некоторое (отличное от нуля) число, вычесть из соответствующих частей другого уравнения и составить систему, в которой вместо одного из упомянутых уравнений стоит уравнение, полученное в результате вычитания, а остальные уравнения оставлены без изменений, то полученная система будет эквивалентна исходной.

Например, каждая из систем

$$\begin{aligned} a_{11}x+\quad &a_{12}y+\quad a_{13}z=b_1,\\ (a_{11}-ca_{21})x+(a_{12}-&ca_{22})y+(a_{13}-ca_{23})z=b_1-cb_2,\\ a_{31}x+\quad &a_{32}y+\quad a_{33}z=b_3; \end{aligned}$$

$$\begin{aligned} (a_{11}-ca_{21})x+(a_{12}-&ca_{22})y+(a_{13}-ca_{23})z=b_1-cb_2,\\ a_{21}x+\quad &a_{22}y+\quad a_{23}z=b_2,\\ a_{31}x+\quad &a_{32}y+\quad a_{33}z=b_3 \end{aligned}$$

при $c\neq 0$ будет эквивалентна исходной системе (2) и обе эти системы эквивалентны между собой.

Заметим, что если после выполнения этих преобразований в системе появляется уравнение, у которого все коэффициенты в левой части равны нулю, а правая часть (свободный член) также равна нулю, то очевидно, что это уравнение обращается в тождество при любых значениях неизвестных и после отбрасывания этого уравнения получается система уравнений, эквивалентная исходной системе. Если же свободный член такого уравнения не равен нулю, то это уравнение не может обратиться в тождество ни при каком наборе значений неизвестных, и поэтому полученная система уравнений, равно как и эквивалентная ей исходная система, будет несовместной.

Метод последовательного исключения неизвестных. Пусть дана система s линейных уравнений с n неизвестными

$$\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots&\\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n&=b_s. \end{aligned} \tag{3}$$

Предположим для определенности, что коэффициент a_{11} не равен нулю (хотя он может оказаться и равным нулю, и тогда необходимо

начать описанную ниже процедуру с другого уравнения системы, а именно с уравнения, у которого коэффициент при неизвестном x_1 отличен от нуля). Преобразуем систему (3), исключая неизвестное x_1 из всех уравнений, кроме первого. Для этого обе части первого уравнения умножим на число a_{21}/a_{11} и вычтем из соответствующих частей второго уравнения, затем обе части первого уравнения, умноженные на число a_{31}/a_{11}, вычтем из соответствующих частей третьего уравнения и т. д.

В результате указанных преобразований приходим к системе из s линейных уравнений с n неизвестными

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= b_2^{(1)}, \\ a_{32}^{(1)}x_2 + a_{33}^{(1)}x_3 + \ldots + a_{3n}^{(1)}x_n &= b_3^{(1)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{s2}^{(1)}x_2 + a_{s3}^{(1)}x_3 + \ldots + a_{sn}^{(1)}x_n &= b_s^{(1)}, \end{aligned} \tag{4}$$

где $a_{ij}^{(1)}$ и $b_i^{(1)}$ — новые коэффициенты при неизвестных и новые свободные члены, выражающиеся через коэффициенты и свободные члены исходной системы (3). Система (4) эквивалентна системе (3). Преобразуем теперь систему (4). При этом первое уравнение не будем трогать совсем и будем преобразовывать лишь часть системы (4), состоящую из всех уравнений кроме первого. Кроме того, мы будем считать, что среди этих уравнений нет таких, у которых все коэффициенты левых частей равны нулю,— такие уравнения мы отбросили бы, если бы их свободные члены были равны нулю (в противном случае мы уже доказали бы несовместность нашей системы). Итак, среди коэффициентов $a_{ij}^{(1)}$ имеются отличные от нуля; для определенности примем, что $a_{22}^{(1)} \neq 0$. Преобразуем теперь систему (4), вычитая из обеих частей третьего и каждого из следующих уравнений обе части второго уравнения, умноженные соответственно на числа

$$\frac{a_{32}^{(1)}}{a_{22}^{(1)}}, \quad \frac{a_{42}^{(1)}}{a_{22}^{(1)}}, \quad \ldots, \quad \frac{a_{s2}^{(1)}}{a_{22}^{(1)}}.$$

В результате из этих уравнений, кроме первого и второго, будет исключено неизвестное x_2 и мы придем к следующей системе уравнений, эквивалентной системе (4), а следовательно, и системе (3):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= b_2^{(1)}, \\ a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= b_3^{(2)}, \\ \ldots\ldots\ldots\ldots & \\ a_{t3}^{(2)}x_3 + \ldots + a_{tn}^{(2)}x_n &= b_t^{(2)}. \end{aligned}$$

Система содержит теперь t уравнений, причем $t \leqslant s$, так как некоторые уравнения оказались, возможно, отброшенными. В дальнейшем будем подвергать преобразованиям лишь часть полученной системы, содержащую все уравнения кроме двух первых.

В результате этого процесса последовательного исключения неизвестных может получиться, что:

1) либо мы придем к системе, в которой одно из уравнений имеет отличный от нуля свободный член, а все коэффициенты левой части равны нулю. В этом случае исходная система несовместна;

2) либо мы получим следующую совместную систему уравнений, эквивалентную системе (3):

$$\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3+\ldots+\quad a_{1n}x_n&=b_1,\\ a_{22}^{(1)}x_2+a_{23}^{(1)}x_3+\ldots+\quad a_{2n}^{(1)}x_n&=b_2^{(1)},\\ a_{33}^{(2)}x_3+\ldots+\quad a_{3n}^{(2)}x_n&=b_3^{(2)},\\ \ldots\ldots\ldots\ldots\ldots\ldots&\\ a_{kk}^{(p-1)}x_k+\ldots+a_{kn}^{(p-1)}x_n&=b_k^{(p-1)}. \end{aligned} \tag{5}$$

Здесь $a_{11}\neq 0$, $a_{22}^{(1)}\neq 0$, $a_{33}^{(2)}\neq 0$, $\ldots$, $a_{kk}^{(p-1)}\neq 0$, а верхний индекс, стоящий в скобках, указывает на номер преобразования, после которого получены эти коэффициенты при неизвестных и новые свободные члены. Отметим также, что $k\leqslant s$ и, очевидно, $k\leqslant n$.

Эта система будет определенной при $k=n$ (число уравнений равно числу неизвестных) и неопределенной при $k<n$ (число уравнений меньше числа неизвестных). Действительно, при $k=n$ система (5) имеет вид

$$\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3+\ldots+a_{1n}x_n&=b_1,\\ a_{22}^{(1)}x_2+a_{23}^{(1)}x_3+\ldots+a_{2n}^{(1)}x_n&=b_2^{(1)},\\ a_{33}^{(2)}x_3+\ldots+a_{3n}^{(2)}x_n&=b_3^{(2)},\\ \ldots\ldots\ldots\ldots\ldots\ldots&\\ a_{nn}^{(p-1)}x_n&=b_n^{(p-1)}, \end{aligned} \tag{6}$$

поэтому из последнего уравнения мы получаем вполне определенное значение для x_n. Подставив его в предпоследнее уравнение, найдем единственное значение для x_{n-1}.

Подставляя найденные значения x_n и x_{n-1} в третье снизу уравнение системы, находим значение x_{n-2}. Продолжая эту процедуру подстановки найденных значений неизвестных в оставшиеся уравнения, придем к тому, что система (6), а следовательно, и система (3) имеют единственное решение, т. е. обе системы являются совместными и определенными (в этом случае также говорят, что система (3) приводится к *треугольному виду*).

Если же $k<n$, то для «свободных» неизвестных x_{k+1}, $\ldots$, x_n возьмем произвольные числовые значения, после чего, двигаясь по системе (5) снизу вверх, как и выше, найдем для неизвестных x_k, x_{k-1}, $\ldots$, x_2, x_1 вполне определенные значения. Поскольку значения свободных неизвестных можно выбирать бесконечным числом различных способов, то система (5), а следовательно, и система (3) будут совместными, но неопределенными. (В этом случае говорят, что система приводится к *трапецеидальной форме*.)

Итак, метод Гаусса применим к любой системе линейных уравнений. При этом система будет несовместной, если в процессе преобразований будет получено уравнение, в котором коэффициенты при всех неизвестных равны нулю, а свободный член отличен от нуля; если же такое уравнение не встретится, то система будет совместной. Совместная система уравнений будет определенной, если она приводится к треугольному виду (6), и неопределенной, если она приводится к трапецеидальному виду (5) при $k<n$.

Пример. Исследовать, при каких значениях параметров a и b система трех линейных уравнений с тремя неизвестными

$$\begin{aligned} x+y+z&=0,\\ 2x+3y+z&=1,\\ x+2y+az&=b \end{aligned} \tag{7}$$

а) имеет единственное решение (найти это решение); б) несовместна; в) неопределенна.

Будем исследовать систему (7) методом Гаусса. Умножая обе части второго уравнения системы на $-1/2$, а обе части третьего уравнения на -1, получим систему, эквивалентную системе (7):

$$\begin{aligned} x+y+z&=0,\\ -x-\frac{3}{2}y-\frac{1}{2}z&=-\frac{1}{2},\\ -x-2y-az&=-b. \end{aligned} \tag{8}$$

Сложим первое уравнение системы (8) со вторым, а третье — с первым; уравнения, полученные в результате сложения, возьмем в качестве второго и третьего уравнений новой системы, эквивалентной системе (8):

$$\begin{aligned} x+y+z&=0,\\ -\frac{1}{2}y+\frac{1}{2}z&=-\frac{1}{2},\\ -y+(1-a)z&=-b. \end{aligned} \tag{9}$$

Теперь умножим обе части третьего уравнения на $-1/2$ и сложим второе уравнение с преобразованным третьим уравнением; полученное уравнение возьмем в качестве третьего уравнения системы. В результате этих преобразований система (9) примет вид

$$\begin{aligned} x+y+z&=0,\\ -\frac{1}{2}y+\frac{1}{2}z&=-\frac{1}{2},\\ \frac{1}{2}az&=\frac{1}{2}b-\frac{1}{2}. \end{aligned}$$

Умножая обе части второго уравнения на (-2), а третьего — на 2, окончательно получаем систему

$$\begin{aligned} x+y+z&=0,\\ y-z&=1,\\ az&=b-1. \end{aligned} \tag{10}$$

В зависимости от значений параметров a и b возможны следующие случаи:

1) при $a\neq 0$ и любом b система имеет единственное решение

$$x=\frac{2-2b-a}{a},\quad y=\frac{a+b-1}{a},\quad z=\frac{b-1}{a};$$

2) при $a=0$ и $b\neq 1$ левая часть третьего уравнения системы (10) обращается в нуль, а правая отлична от нуля. Следовательно, при этих значениях параметров a и b система оказывается несовместной;

3) при $a=0$ и $b=1$ система (10) принимает трапецеидальную форму. При этих значениях параметров a и b система совместна, неопределенна и имеет вид

$$\begin{aligned} x+y+z&=0,\\ y-z&=1. \end{aligned}$$

Полагая $z=c$, где c — любое число, получаем множество решений системы:

$$x=-2c-1,\quad y=c+1,\quad z=c.$$

Метод Гаусса решения систем линейных уравнений с числовыми коэффициентами в силу простоты и однотипности выполняемых операций пригоден для счета на электронно-вычислительных машинах.

Существенным недостатком этого метода является невозможность сформулировать условия совместности и определенности системы в зависимости от значений коэффициентов и свободных членов. С другой стороны, даже в случае определенной системы этот метод не позволяет найти общие формулы, выражающие решение системы через ее коэффициенты и свободные члены, которые необходимо иметь при теоретических исследованиях. Существуют и другие методы решения и исследования систем линейных уравнений, которые лишены отмеченных недостатков. Эти методы основаны на теории матриц и определителей (см. § 8).

7.4. Некоторые способы решения систем нелинейных алгебраических уравнений. Для систем нелинейных алгебраических уравнений, в отличие от систем линейных уравнений, не найдено какого-либо универсального практически удобного метода решения, а потому при решении каждой конкретной системы нелинейных уравнений приходится применять специальные приемы решений, основанные на использовании особенностей алгебраических уравнений, составляющих данную систему.

Разберем на примерах два приема решения систем нелинейных алгебраических уравнений.

Пример 1. Решить систему уравнений

$$\begin{aligned} 2x^2-3xy+y^2&=0,\\ 9x^2-2xy-y^2+5x-5y&=6. \end{aligned} \tag{1}$$

Замечая, что многочлен двух переменных в левой части первого уравнения системы (1) однородный, систему (1) можно свести к двум системам, эквивалентным системе (1):

$$\begin{aligned} 2x^2-3xy+y^2&=0,\\ y&=0,\\ 9x^2-2xy-y^2+5x-5y&=6; \end{aligned} \tag{2}$$

$$\begin{aligned} 2\left(\frac{x}{y}\right)^2-3\left(\frac{x}{y}\right)+1&=0,\\ 9x^2-2xy-y^2+5x-5x&=6. \end{aligned} \tag{2'}$$

Найдем решения системы (2). Подставляя значение $y=0$ в первое и третье уравнения системы (2), получаем систему двух уравнений для нахождения значения одного неизвестного x:

$$x=0, \quad 9x^2+5x=6.$$

Поскольку не существует числа x, удовлетворяющего обоим уравнениям системы, то система (2) несовместна.

Перейдем к решению системы (2'). Решим первое уравнение системы как квадратное относительно переменной $u=x/y$. Уравнение

$$2u^2-3u+1=0$$

имеет корни $u_1=1$, $u_2=1/2$, т. е. $x/y=1$, $x/y=1/2$.

Таким образом, система (2') эквивалентна двум системам:

$$\begin{aligned} y&=x,\\ 9x^2-2xy-y^2+5x-5y&=6; \end{aligned} \tag{3}$$

$$\begin{aligned} y&=2x,\\ 9x^2-2xy-y^2+5x-5y&=6. \end{aligned} \tag{3'}$$

Подставляя выражения y через x во вторые уравнения систем, получаем соответствующие им уравнения для нахождения x:

$$x^2=1, \quad x^2-5x-6=0.$$

Решения первого уравнения: $x_1=1$, $x_2=-1$.

Решения второго уравнения: $x_3=6$, $x_4=-1$.

Системы (3) и (3') имеют соответственно решения

$$\begin{aligned} x_1&=1, \quad y_1=1; \quad & x_2&=-1, \quad y_2=-1;\\ x_3&=6; \quad y_3=12; \quad & x_4&=-1, \quad y_4=-2, \end{aligned}$$

которые и будут решениями системы (1).

Симметрические системы уравнений. Система уравнений с двумя неизвестными x и y называется *симметрической*, если она не меняется при замене неизвестного x на неизвестное y, а неизвестного y на неизвестное x.

Часто решение таких систем может быть найдено с помощью введения новых переменных — элементарных *симметрических* многочленов σ_1 и σ_2:

$$\sigma_1 = x + y, \quad \sigma_2 = xy.$$

Пример 2. Найти решения системы

$$\begin{aligned} x^2 + xy + y^2 &= 4, \\ x + xy + y &= 2. \end{aligned} \tag{4}$$

Введем новые переменные — элементарные симметрические многочлены $\sigma_1 = x + y$, $\sigma_2 = xy$. Относительно переменных σ_1 и σ_2 система (4) запишется в виде

$$\begin{aligned} \sigma_1^2 - \sigma_2 &= 4, \\ \sigma_1 + \sigma_2 &= 2. \end{aligned} \tag{5}$$

Выражая σ_2 из второго уравнения системы через σ_1:

$$\sigma_2 = 2 - \sigma_1, \tag{6}$$

и подставляя $\sigma_2 = 2 - \sigma_1$ в первое уравнение системы (5), получаем квадратное уравнение $\sigma_1^2 + \sigma_1 - 6 = 0$, корни которого равны 2 и -3.

Подставляя найденные значения σ_1 в равенство (6), получаем значения σ_2. Таким образом, множество решений системы (5) имеет вид

$$\sigma_1 = 2, \quad \sigma_2 = 0; \qquad \sigma_1 = -3, \quad \sigma_2 = 5.$$

Теперь множество решений исходной системы (4) может быть получено как объединение множеств решений двух более простых систем:

$$\begin{aligned} x + y &= 2, & x + y &= -3, \\ xy &= 0; & xy &= 5. \end{aligned}$$

Вторая система решений не имеет. Решениями первой системы являются пары чисел $x_1 = 2$, $y_1 = 0$ и $x_2 = 0$, $y_2 = 2$.

Пример 3. Решить систему уравнений

$$\begin{aligned} x + y &= 5, \\ x^6 + y^6 &= 275. \end{aligned}$$

Замена $\sigma_1 = x + y$, $\sigma_2 = xy$ приводит данную систему уравнений к виду

$$\begin{aligned} \sigma_1 &= 5, \\ \sigma_1^6 - 5\sigma_1^3\sigma_2 + 5\sigma_1\sigma_2^2 &= 275. \end{aligned}$$

Подставляя значение $\sigma_1 = 5$ во второе уравнение, получим уравнение для σ_2: $\sigma_2^2 - 25\sigma_2 + 114 = 0$. Корнями этого уравнения будут числа 6 и 19.

Таким образом, множество решений исходной системы находится как объединение множеств решений двух более простых систем:

$$\begin{aligned} x + y &= 5, & x + y &= 5, \\ xy &= 6; & xy &= 19. \end{aligned}$$

Решим первую систему. Из первого уравнения имеем $y = 5 - x$. Подставляя $y = 5 - x$ во второе уравнение, получаем квадратное уравнение относительно неизвестного x: $x^2 - 5x + 6 = 0$, корнями которого являются $x_1 = 3$ и $x_2 = 2$. Из равенства $y = 5 - x$ находим $y_1 = 5 - x_1 = 2$ и $y_2 = 5 - x_2 = 3$. Итак, решения первой системы — пары чисел $x_1 = 3$, $y_1 = 2$ и $x_2 = 2$, $y_2 = 3$.

Аналогично можно убедиться в том, что вторая система не имеет решений.

Симметрическая система трех уравнений с тремя неизвестными x, y, z обычно решается с помощью введения новых неизвестных — элементарных симметрических многочленов

$$\sigma_1 = x + y + z, \quad \sigma_2 = xy + yz + xz, \quad \sigma_3 = xyz.$$

В п. 7.3 были рассмотрены системы линейных алгебраических уравнений, в которых число уравнений было меньше числа неизвестных. При этом оказывалось, что одни неизвестные выражались через другие («свободные») неизвестные, т. е. система имела бесчисленное множество решений. По-другому обстоит дело в случае решения систем нелинейных алгебраических уравнений, в которых число неизвестных больше числа уравнений. Может оказаться, что такая система имеет как бесконечное, так и конечное число решений или вовсе не имеет решений. Сейчас мы рассмотрим в качестве примера систему двух алгебраических уравнений с тремя неизвестными, которая имеет единственное решение (т. е. множество решений представляет собой единственную упорядоченную тройку чисел, обращающую оба уравнения в тождества).

Пример 4. Найти действительные решения системы

$$\begin{aligned} x+y&=2,\\ xy-z^2&=1. \end{aligned} \tag{7}$$

Выражая y через x из первого уравнения и подставляя $y=2-x$ во второе, получаем уравнение $(x-1)^2+z^2=0$, которое эквивалентно системе

$$x-1=0, \quad z=0.$$

Эта система имеет единственное решение, поэтому и система (7) двух уравнений с тремя неизвестными имеет только одно решение (1; 1; 0).

§ 8. Исследование систем линейных уравнений с помощью определителей

8.1. Матрицы и операции над ними. Прямоугольная таблица

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{Vmatrix} \tag{1}$$

из $m \cdot n$ чисел называется *матрицей* из m строк и n столбцов или *матрицей размера* $m \times n$, а также $m \times n$-*матрицей*.

Числа a_{ij} ($i=1, 2, \dots, m$; $j=1, 2, \dots, n$) называются *элементами* матрицы; первый индекс i указывает номер строки, в которой стоит элемент матрицы, а второй индекс j — номер столбца.

Матрица (1) может обозначаться также

$$\|a_{ij}\|, \quad i=1, 2, \dots, m; \qquad i=1, 2, \dots, n.$$

Кроме того, для матриц используются обозначения

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \quad \text{или} \quad (a_{ij});$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \quad \text{или} \quad [a_{ij}].$$

Если число строк матрицы равно числу столбцов (и равно n), то матрица называется *квадратной матрицей* (*порядка* n).

Две матрицы $\|a_{ij}\|$ и $\|b_{kl}\|$ называются *равными*, если они имеют равное число строк и столбцов соответственно и если равны числа,

стоящие на соответственных местах

$$a_{ij} = b_{kl} \quad \text{при} \quad i = k \quad \text{и} \quad j = l.$$

Основными арифметическими операциями над матрицами являются умножение матрицы на число, сложение и умножение матриц. Далее для краткости матрицы также будем обозначать буквами A, B, ...

Умножение матрицы на число. *Произведением* числа λ и матрицы $A = \|a_{ij}\|$ называется матрица $B = \|b_{ij}\|$, элементы которой вычисляются по правилу $b_{ij} = \lambda \cdot a_{ij}$, т. е. каждый элемент матрицы b_{ij} представляет собой произведение числа λ и элемента матрицы a_{ij}.

Например,

$$\lambda \cdot \begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix} = \begin{Vmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{Vmatrix}.$$

Матрица, все элементы которой равны нулю, называется *нулевой* матрицей и обычно обозначается символом 0 (или 0_{ij}).

Свойства операции умножения матрицы на число:

1) коммутативность:

$$\lambda \cdot A = A \cdot \lambda;$$

2) ассоциативность:

$$(\alpha \cdot \beta) \cdot A = \alpha \cdot (\beta \cdot A).$$

Сложение матриц. *Суммой* двух матриц $A = \|a_{ij}\|$ и $B = \|b_{kl}\|$, имеющих соответственно равные числа строк и столбцов (i, $k = 1$, 2, ..., m и j, $l = 1$, 2, ..., n), называется матрица $S = \|s_{ij}\|$ размера $m \times n$ с элементами, равными суммам соответственных элементов матриц A и B: $s_{ij} = a_{ij} + b_{ij}$. Сумма матриц A и B обозначается $A + B$.

Например, суммой матриц

$$\begin{Vmatrix} 2 & 1 & 1 \\ 0 & -1 & 5 \end{Vmatrix} \quad \text{и} \quad \begin{Vmatrix} -2 & -2 & 3 \\ 2 & 3 & 1 \end{Vmatrix}$$

является матрица

$$\begin{Vmatrix} 0 & -1 & 4 \\ 2 & 2 & 6 \end{Vmatrix}.$$

Свойства операции сложения матриц:

1) коммутативность:

$$A + B = B + A;$$

2) ассоциативность:

$$(A + B) + C = A + (B + C);$$

3) дистрибутивность относительно сложения матриц:

$$\lambda \cdot (A + B) = \lambda \cdot A + \lambda \cdot B;$$

4) дистрибутивность относительно сложения чисел:

$$(\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A.$$

Перечисленные свойства операций умножения матрицы на число и операции сложения матриц есть следствия определения этих операций и свойств коммутативности и ассоциативности операций сложения и умножения действительных чисел, а также дистрибутивности умножения относительно сложения.

Операция сложения матриц имеет обратную операцию — вычитание. *Разностью* матриц A и B называется матрица C, составленная

из разностей соответственных элементов заданных матриц A и B. О матрице C говорят, что она получена в результате вычитания матрицы B из матрицы A, и пишут $C=A-B$.

Умножение матриц. Пусть A и B — две матрицы, причем число столбцов первой матрицы равно числу строк второй матрицы, т. е. эти матрицы имеют вид

$$A=\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix} \quad \text{и} \quad B=\begin{Vmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{Vmatrix}.$$

Произведением матриц $A=\|a_{ij}\|$ и $B=\|b_{ij}\|$ называется матрица C порядка $m\times k$, элементы которой вычисляются по правилу

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\dots+a_{in}b_{nj}, \quad i=1, 2, \dots, m, \quad j=1, 2, \dots, k,$$

т. е. элемент, стоящий в i-й строке и j-м столбце произведения двух матриц, получается в результате умножения первого элемента i-й строки матрицы A на первый элемент j-го столбца матрицы B, второго элемента i-й строки матрицы A на второй элемент j-го столбца матрицы B и т. д. и последующего сложения всех таких произведений пар элементов матриц A и B. Произведение матриц A и B обозначается $A\cdot B$. Согласно данному определению произведение матриц $A\cdot B$ существует, если число строк матрицы A равно числу столбцов матрицы B.

Операция умножения матриц, вообще говоря, не коммутативна. Например,

$$\begin{Vmatrix} 1 & 2 \\ 3 & 4 \end{Vmatrix}\cdot\begin{Vmatrix} 1 & 1 \\ 0 & 1 \end{Vmatrix}=\begin{Vmatrix} 1 & 3 \\ 3 & 7 \end{Vmatrix}; \qquad \begin{Vmatrix} 1 & 1 \\ 0 & 1 \end{Vmatrix}\cdot\begin{Vmatrix} 1 & 2 \\ 3 & 4 \end{Vmatrix}=\begin{Vmatrix} 4 & 6 \\ 3 & 4 \end{Vmatrix},$$

т. е. произведение матриц зависит от порядка следования сомножителей. Более того, если перемножить неквадратные матрицы, то может оказаться, что произведение двух матриц, перемножаемых в одном порядке, существует, а в другом — нет.

Операция умножения матриц ассоциативна:

$$(A\cdot B)\cdot C=A\cdot(B\cdot C).$$

Основные свойства операции умножения матриц*):

$$\alpha\cdot(A\cdot B)=(\alpha\cdot A)\cdot B; \qquad A\cdot(\alpha\cdot B)=(A\cdot\alpha)\cdot B;$$
$$(A\cdot B)\cdot\alpha=A\cdot(B\cdot\alpha); \qquad (A+B)\cdot C=A\cdot C+B\cdot C;$$
$$C\cdot(A+B)=C\cdot A+C\cdot B.$$

В множестве квадратных матриц фиксированного порядка n операции сложения и умножения определены для любых матриц — суммой и произведением двух квадратных матриц порядка n снова будут квадратные матрицы порядка n, т. е. множество всех квадратных матриц фиксированного порядка замкнуто относительно операций сложения и умножения, причем операция умножения ассоциативна, но, вообще говоря, не коммутативна.

Элементы $a_{11}, a_{22}, \dots, a_{nn}$ квадратной матрицы порядка n называются *диагональными элементами*.

*) Эти свойства справедливы, естественно, лишь в том случае, если указанные действия над матрицами имеют смысл (т. е. только для тех матриц, число строк и столбцов которых удовлетворяет соответствующим условиям).

Квадратная матрица, все диагональные элементы которой равны 1, а остальные — нули, называется *единичной* и обозначается E (или E_n, где n — ее порядок, или $\|\delta_{ij}\|$, где $i, j = 1, 2, \ldots, n$. Символ δ_{ij} имеет специальное название — «символ Кронекера»).

Для любой квадратной матрицы A порядка n справедливо равенство

$$E_n \cdot A = A \cdot E_n = A.$$

Матрицы $\|a_{ij}\|$ порядка n, у которых $a_{ij} = 0$ для всех $i \neq j$, т. е. матрицы вида

$$\begin{Vmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22} & 0 & \ldots & 0 \\ 0 & 0 & a_{33} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{nn} \end{Vmatrix}$$

называются *диагональными матрицами* (об элементах $a_{11}, a_{22}, \ldots, a_{nn}$ также говорят, что они стоят на главной диагонали).

Свойство диагональных матриц. Сумма и произведение двух диагональных матриц — также диагональные матрицы.

Транспонирование матриц. Пусть $A = \|a_{ij}\|$ — матрица размера $m \times n$:

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix}.$$

Матрица, получающаяся из матрицы A заменой строк столбцами, называется *транспонированной* матрицей по отношению к матрице A и обозначается $A_т$.

Квадратная матрица A называется *симметрической*, если $A_т = A$, и *кососимметрической*, если $A_т = -A$.

Элементы, расположенные симметрично относительно главной диагонали, у симметрической матрицы равны, а у кососимметрической противоположны. Все диагональные элементы кососимметрической матрицы равны нулю.

8.2. Определители. Пусть $A = \|a_{ij}\|$ — матрица n-го порядка:

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{Vmatrix}.$$

Составим всевозможные произведения по n элементов этой матрицы, расположенных в разных строках и разных столбцах, т. е. произведения вида

$$a_{1i_1} \cdot a_{2i_2} \cdot \ldots \cdot a_{ni_n}, \qquad (1)$$

где индексы $i_1, i_2, \ldots, i_n$ образуют некоторую перестановку из чисел $1, 2, \ldots, n$. Число таких перестановок (а следовательно, и всевозможных произведений вида (1)) равно $n!$

Определителем (или *детерминантом*) квадратной матрицы n-го порядка

$$\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Vmatrix}$$

называется алгебраическая сумма всевозможных различных произведений

$$\sum (-1)^p a_{1i_1} \cdot a_{2i_2} \cdot \dots \cdot a_{ni_n},$$

где $p=1$, если подстановка

$$\begin{pmatrix} 1 & 2 & 3 & \dots & n \\ i_1 & i_2 & i_3 & \dots & i_n \end{pmatrix}$$

нечетная, и $p=2$, если подстановка четная.

Определитель квадратной матрицы n-го порядка $A=\|a_{ij}\|$ ($i, j=1, 2, \dots, n$). обозначается

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \quad \text{или} \quad \det\|a_{ij}\|, \quad \text{или} \quad |a_{ij}|.$$

Примеры. 1. Определителем матрицы второго порядка

$$A=\begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}$$

является выражение

$$a_{11}a_{22}-a_{12}a_{21}.$$

2. Определителем матрицы третьего порядка

$$A=\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}$$

является выражение

$$a_{11}a_{22}a_{23}+a_{13}a_{21}a_{32}+a_{12}a_{23}a_{31}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}.$$

Правило вычисления определителя матрицы третьего порядка графически можно изобразить следующим образом: на рис. 4.2 прямыми линиями соединены элементы матрицы, произведение которых входит в определитель со знаком плюс, а на рис. 4.3 прямыми соединены элементы матрицы, произведение которых входит в определитель со знаком минус.

Основные свойства определителей:

1) определитель транспонированной матрицы равен определителю исходной матрицы;

2) если одна из сторон определителя состоит только из нулей, то определитель равен нулю;

3) при перестановке местами двух строк определителя получается определитель противоположного знака;

4) определитель, имеющий две одинаковые строки, равен нулю;

5) если все элементы некоторой строки определителя умножить на число k, то сам определитель умножится на k;

6) если одна из строк определителя пропорциональна другой строке, то определитель равен нулю;

7) если каждый из элементов i-й строки определителя n-го порядка представить в виде суммы двух слагаемых a_j+b_j ($j=1, 2, \ldots, n$), то определитель будет равен сумме двух определителей, у которых все строки, кроме i-й, такие же, как и в исходном определителе, а i-я строка одного из определителей состоит из элементов a_j, а другого — из элементов b_j;

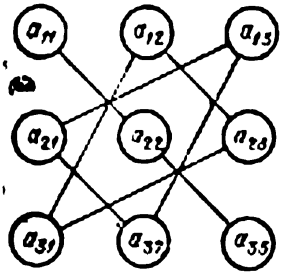


Рис 4.2

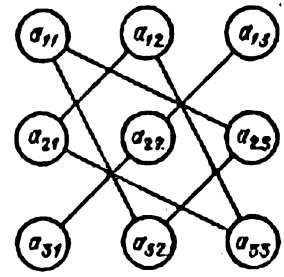


Рис 4.3

8) определитель не меняется, если к элементам одной строки прибавляются элементы другой, умноженные на одно и то же число.

Практически удобный способ вычисления определителей порядка выше третьего основан на следующем свойстве определителей, которое называется *разложением определителя по строке*.

Пусть $\|a_{ij}\|$ — матрица n-го порядка. Определитель этой матрицы может быть записан в виде

$$\det\|a_{ij}\| = a_{11}\cdot|A_{11}| - a_{12}\cdot|A_{12}| + \ldots + (-1)^{n-1}a_{1n}\cdot|A_{1n}|,$$

где через $|A_{1j}|$ обозначен определитель матрицы $(n-1)$-го порядка, полученной из матрицы $\|a_{ij}\|$ вычеркиванием первой строки и j-го столбца.

С помощью этого свойства определителей нетрудно свести вычисление определителя квадратной матрицы, например, четвертого порядка к вычислению определителей третьего порядка, если записать ее определитель в виде

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = a_{11}\cdot\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12}\cdot\begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} +$$

$$+ a_{13}\cdot\begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14}\cdot\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix}.$$

8.3. Ранг матрицы. Пусть $A=\|a_{ij}\|$ — матрица размера $m\times n$ ($i=1, 2, \ldots, m$; $j=1, 2, \ldots, n$), т. е. матрица, состоящая из m строк и n столбцов. Выберем какие-нибудь k различных строк и k различных столбцов матрицы A. Из элементов матрицы A, находящихся на пересечениях отмеченных строк и отмеченных столбцов, записанных

в их естественном порядке, составим квадратную матрицу порядка k. Определитель этой квадратной матрицы порядка k называют *минором* k-го порядка матрицы A.

Рангом матрицы A называют наивысший порядок отличных от нуля миноров матрицы A.

Пример. Найти ранг матрицы

$$A=\begin{Vmatrix} 2 & -4 & 3 \\ 1 & -2 & 1 \\ 0 & 1 & 0 \end{Vmatrix}.$$

На пересечении первой строки и первого столбца стоит элемент матрицы $a_{11}=2$. Следовательно, ранг матрицы A не меньше единицы.

Из элементов, стоящих на пересечении первых двух строк и первых двух столбцов, образуем матрицу

$$\begin{Vmatrix} 2 & -4 \\ 1 & -2 \end{Vmatrix},$$

определитель которой равен нулю. Однако в матрице A содержатся и отличные от нуля миноры второго порядка. Таким минором, например, будет определитель матрицы, составленной из элементов, стоящих на пересечении первого и второго столбцов и второй и третьей строк:

$$\begin{vmatrix} 1 & -2 \\ 0 & 1 \end{vmatrix}=1.$$

Следовательно, ранг данной матрицы по крайней мере равен двум.

Минором следующего (третьего) порядка будет определитель матрицы A, который равен нулю.

Следовательно, ранг данной матрицы равен двум.

Ранг матрицы может быть найден и без вычисления различных миноров этой матрицы. Этот метод вычисления ранга матрицы основан на том, что любая матрица может быть приведена к диагональной форме с помощью следующих элементарных преобразований матрицы, не меняющих ее ранга:

1) перемеой местами двух строк или двух столбцов;

2) умножением строки или столбца на произвольное отличное от нуля число;

3) прибавлением к одной строке (столбцу) другой строки (столбца), умноженной на некоторое число.

Для нахождения ранга матрицы нужно элементарными преобразованиями привести эту матрицу к диагональной форме. Число отличных от нуля элементов, стоящих на главной диагонали полученной матрицы, равно рангу матрицы.

Пример. Найти ранг матрицы

$$A=\begin{Vmatrix} 0 & 2 & -4 \\ -1 & -4 & 5 \\ 3 & 1 & 7 \\ 0 & 5 & -10 \end{Vmatrix}.$$

Меняя местами в этой матрице первый и второй столбцы и умножая новую первую строку на 1/2, получаем матрицу, у которой элемент a_{11} равен 1:

$$\begin{Vmatrix} 1 & 0 & -2 \\ -4 & -1 & 5 \\ 1 & 3 & 7 \\ 5 & 0 & -10 \end{Vmatrix}.$$

Прибавляя к третьему столбцу удвоенный первый столбец, получаем матрицу

$$\begin{Vmatrix} 1 & 0 & 0 \\ -4 & -1 & -3 \\ 1 & 3 & 9 \\ 5 & 0 & 0 \end{Vmatrix}.$$

Преобразуем строки матрицы, начиная со второй. Умножим первую строку на 4 и прибавим ко второй строке; умножим первую строку на (—1) и прибавим к третьей строке; умножим первую строку на (—5) и прибавим к четвертой строке. В результате этих элементарных преобразований получится матрица

$$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & -1 & -3 \\ 0 & 3 & 9 \\ 0 & 0 & 0 \end{Vmatrix}.$$

Далее последовательно будем преобразовывать строки и столбцы матрицы, не меняя первой строки и первого столбца:

1) умножим вторую строку на (—1);

2) умножим вторую строку на (—3) и прибавим к третьей строке;

3) умножим второй столбец на (—3) и прибавим к третьему столбцу. В результате этих трех элементарных преобразований последовательно получаются матрицы

$$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 3 & 9 \\ 0 & 0 & 0 \end{Vmatrix}, \quad \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}, \quad \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}.$$

Последняя матрица имеет диагональную форму; число единиц, стоящих на главной диагонали, равно двум. Ранг этой матрицы (а следовательно, и ранг матрицы A) равен двум.

8.4. Исследование систем линейных уравнений с помощью определителей. Формулы Крамера. Существенный недостаток решения систем линейных уравнений методом последовательного исключения неизвестных (см. п. 7.3) состоит в том, что он не позволяет сформулировать условия совместности и определенности системы линейных уравнений через ее коэффициенты и свободные члены, а также (в случае определенной системы) не дает формул, выражающих решение системы через ее коэффициенты и свободные члены, которые могут быть полезными при теоретических исследованиях.

Общие формулы, выражающие решение системы линейных уравнений через ее коэффициенты и свободные члены, а также условия совместности и определенности системы линейных уравнений основаны на использовании понятий определителя матрицы (см. п. 8.2) и ранга матрицы (см. п. 8.3). Рассмотрим систему m линейных уравнений с n неизвестными $x_1, x_2, \ldots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m. \end{aligned} \tag{2}$$

Для записи системы уравнений (2) используется и более короткая (матричная) форма записи. Записывая неизвестные $x_1, x_2, \ldots, x_n$ и свободные члены $b_1, b_2, \ldots, b_m$ в виде столбцов (матриц размера

$n\times 1$ и $m\times 1$ соответственно) и обозначая эти столбцы через X и B:

$$\left\|\begin{matrix} x_1 \\ x_2 \\ \dots \\ x_n \end{matrix}\right\| = X, \qquad \left\|\begin{matrix} b_1 \\ b_2 \\ \dots \\ b_m \end{matrix}\right\| = B,$$

систему (2) можно записать в виде

$$A \cdot X = B.$$

Матрица

$$A = \|a_{ij}\| = \left\|\begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{matrix}\right\|$$

называется *основной* матрицей системы (2), а матрица

$$\overline{A} = \left\|\begin{matrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{matrix}\right\|$$

— *расширенной* матрицей этой системы, причем ранг матрицы $\overline{A}$ либо равен рангу матрицы A, либо на единицу превышает его.

Определитель матрицы A называется *определителем системы* (2).

Система уравнения называется *крамеровской*, если число уравнений равно числу неизвестных и определитель системы отличен от нуля.

Любая крамеровская система уравнений имеет единственное решение $(x_1, x_2, \dots, x_n)$, даваемое формулами

$$x_i = \frac{\Delta_{x_i}}{\Delta}, \tag{3}$$

где Δ_{x_i} — определитель матрицы, полученной из матрицы A заменой i-го столбца на столбец свободных членов системы (2). Формулы (3) называются *формулами Крамера*.

Полный ответ на вопрос о существовании решения системы m линейных уравнений с n неизвестными дает следующая теорема Кронекера — Капелли.

Для того чтобы система уравнений (2) *имела решение, необходимо и достаточно, чтобы ранг расширенной матрицы был равен рангу основной матрицы системы. Если ранги основной и расширенной матриц совпадают с числам неизвестных, то система имеет единственное решение. Если ранг r основной и расширенной матриц меньше числа неизвестных $(r < n)$, то система* (2) *имеет более одного решения.*

В последнем случае число «свободных» неизвестных, через которые выражаются остальные неизвестные, равно $n - r$.

Система двух линейных уравнений с двумя неизвестными. Система двух линейных уравнений с двумя неизвестными

$$\begin{aligned} a_{11}x + a_{12}y &= b_1, \\ a_{21}x + a_{22}y &= b_2 \end{aligned} \tag{4}$$

имеет следующие основную и расширенную матрицы:

$$A=\begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}, \quad \overline{A}=\begin{Vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{Vmatrix}.$$

причем ранг матрицы $\overline{A}$ равен рангу матрицы A или на единицу больше его.

Умножим левую и правую части первого уравнения системы (4) на a_{22}, второго — на $(-a_{12})$ и сложим эти уравнения. В результате получим уравнение

$$(a_{11}a_{22}-a_{12}a_{21})\cdot x=b_1a_{22}-b_2a_{12}. \tag{5}$$

Аналогично, умножая обе части первого уравнения на $(-a_{21})$, второго — на a_{11} и складывая эти уравнения, получим уравнение

$$(a_{11}a_{22}-a_{12}a_{21})\cdot y=a_{11}b_2-a_{21}b_1. \tag{5'}$$

Коэффициент при неизвестных x и y в левых частях уравнений (5) и (5′) один и тот же и представляет собой определитель матрицы A, составленной из коэффициентов системы уравнений (4)

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

В первых частях уравнений (5) и (5′) также стоят определители матриц второго порядка: правая часть уравнения (5) является определителем матрицы, полученной из матрицы A заменой ее первого столбца на столбец свободных членов системы (4), правая часть равенства (5′) — определитель матрицы, полученной из матрицы A заменой ее второго столбца на столбец свободных членов системы (4).

Уравнения (5), (5′) теперь можно записать в виде

$$\Delta\cdot x=\Delta_x, \qquad \Delta\cdot y=\Delta_y, \tag{6}$$

где через Δ, Δ_x и Δ_y обозначены определители

$$\Delta=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21},$$

$$\Delta_x=\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}=b_1a_{22}-b_2a_{12},$$

$$\Delta_y=\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}=a_{11}b_2-a_{21}b_1.$$

Для системы (6) (а следовательно, и для системы (4)) возможны следующие случаи.

1. Если определитель этой системы

$$\Delta=|A|=\det\begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}$$

не равен нулю, то ранг матрицы A равен двум, ранг матрицы $\overline{A}$ также равен двум, система (4) крамеровская и имеет единственное решение, определяемое формулами Крамера:

$$x=\frac{\Delta_x}{\Delta}, \quad y=\frac{\Delta_y}{\Delta}.$$

2. Если определитель этой системы равен нулю ($\Delta=0$), то ранг матрицы A равен единице (так как хотя бы одно из чисел a_{11}, a_{12}, a_{21}, a_{22} отлично от нуля) и возможны два случая:

а) если ранг расширенной матрицы $\overline{A}$ равен двум, т. е. отличен от нуля хотя бы один из определителей

$$\Delta_x = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \qquad \Delta_y = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix};$$

то согласно теореме Кронекера — Капелли система (4) решений не имеет;

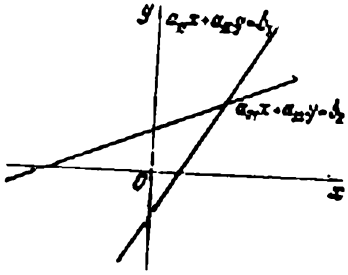


Рис. 4.4

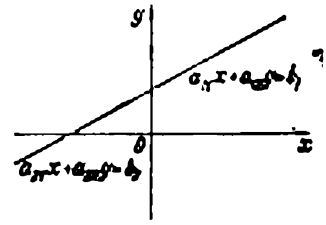


Рис. 4.5

б) если ранг расширенной матрицы $\overline{A}$ равен единице, т. е.

$$\Delta_x = \Delta_y = 0,$$

то согласно теореме Кронекера — Капелли система имеет более одного решения; число свободных неизвестных равно числу неизвестных минус ранг основной матрицы системы, т. е. равно единице, и связь между неизвестными дается одним из уравнений системы.

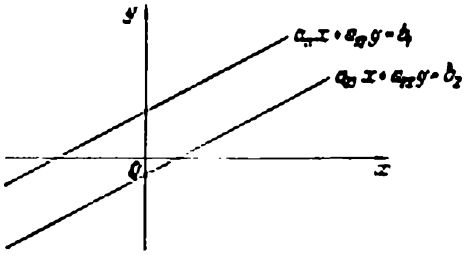


Рис. 4.6

Решения системы двух линейных уравнений с двумя неизвестными имеют следующую геометрическую интерпретацию.

Каждое из уравнений системы (4) задает линейное соответствие между переменными x и y. Всякое линейное соответствие между переменными x и y определяет в прямоугольной системе координат Oxy некоторую прямую. В случае, когда система имеет единственное решение, прямые, задаваемые первым и вторым уравнениями, пересекаются (рис. 4.4). Если система имеет бесчисленное множество решений, то прямые совпадают (рис. 4.5); если система несовместна, то прямые параллельны (рис. 4.6).

Система трех линейных уравнений с тремя неизвестными. Система трех линейных уравнений с тремя неизвестными x, y, z

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1, \\ a_{21}x + a_{22}y + a_{23}z &= b_2, \\ a_{31}x + a_{32}y + a_{33}z &= b_3 \end{aligned} \qquad (7)$$

имеет основную и расширенную матрицы

$$A=\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}, \quad \overline{A}=\begin{Vmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{Vmatrix}.$$

причем ранг матрицы $\overline{A}$ равен рангу матрицы A или на единицу больше его.

Так же как и в случае системы двух линейных уравнений с двумя неизвестными, для неизвестной x можно получить уравнение, коэффициенты которого выражены через коэффициенты системы и ее свободные члены:

$$(a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32})\cdot x = \\ = b_1a_{22}a_{33}+a_{12}a_{23}b_3+a_{13}b_2a_{32}-a_{13}a_{22}b_3-a_{12}b_2a_{33}-b_1a_{23}a_{32}. \quad (8)$$

Коэффициент при x в уравнении (8) представляет собой определитель матрицы A, составленной из коэффициентов системы уравнений (7):

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Правая часть равенства (8) также будет определителем третьего порядка матрицы, полученной из матрицы A заменой ее первого столбца на столбец свободных членов системы (7). Если обозначить определитель матрицы A через Δ, а определитель, стоящий в правой части равенства (8),— через Δ_x,

$$\Delta_x=\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix},$$

то уравнение (8) можно записать в виде

$$\Delta\cdot x=\Delta_x.$$

Аналогично из системы (7) можно получить уравнения для нахождения неизвестных y и z:

$$\Delta\cdot y=\Delta_y, \qquad \Delta\cdot z=\Delta_z,$$

где через Δ_y и Δ_z соответственно обозначены определители третьего порядка матриц, полученных из матрицы A заменой ее второго (соответственно третьего) столбца столбцом свободных членов,

$$\Delta_y=\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \qquad \Delta_z=\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.$$

и система уравнений (7) записывается в виде

$$\Delta\cdot x=\Delta_x, \qquad \Delta\cdot y=\Delta_y, \qquad \Delta\cdot z=\Delta_z. \quad (9)$$

Для системы (9), а следовательно, и для системы (7) возможны следующие случаи:

1) если определитель этой системы

$$\Delta=|A|=\det\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}$$

не равен нулю, то система (7) крамеровская и имеет единственное решение, определяемое формулами Крамера:

$$x=\frac{\Delta_x}{\Delta},\quad y=\frac{\Delta_y}{\Delta},\quad z=\frac{\Delta_z}{\Delta};$$

2) если определитель $\Delta=0$, т. е. ранг матрицы A меньше трех, а ранг расширенной матрицы $\overline{A}$ больше ранга матрицы A, то система не имеет решения (несовместна);

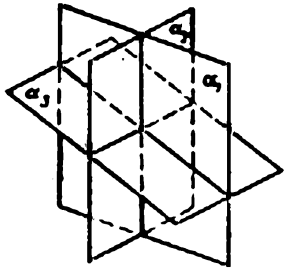


Рис. 4.7

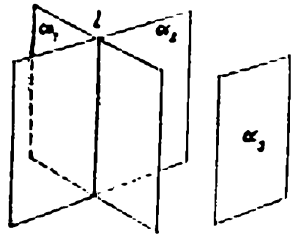


Рис. 4.8

3) если определитель $\Delta=0$ и ранг матрицы A равен рангу расширенной матрицы $\overline{A}$, то система имеет более одного решения. При этом если ранги матрицы A и $\overline{A}$ равны двум (единице), то число свободных неизвестных равно разности числа неизвестных и рангу матриц A и $\overline{A}$, т. е. равно одному (двум).

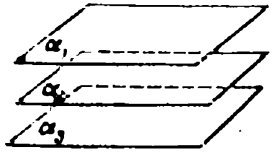


Рис. 4.9

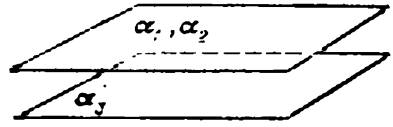


Рис. 4.10

Более подробный анализ случаев 2)—3) приведен ниже, где разобраны все возможные различные соотношения между рангами матриц A и $\overline{A}$.

Линейное уравнение с тремя неизвестными x, y, z определяет плоскость в пространстве и соответственно три уравнения системы (7) — три плоскости. Возможны следующие случаи взаимного расположения этих трех плоскостей α_1, α_2, α_3, задаваемых соответственно первым, вторым и третьим уравнениями.

1. Если система имеет единственное решение (ранг матрицы равен трем), то все три плоскости различны и пересекаются в одной точке (рис. 4.7). Решением системы являются координаты этой точки.

2. Если ранг матрицы A равен двум, а ранг матрицы $\overline{A}$ — трем, то система не имеет решения. Две плоскости пересекаются по прямой l, параллельной третьей плоскости (на рис. 4.8 плоскости α_3).

3. Если ранг матрицы A равен единице, а ранг матрицы $\overline{A}$ — двум, то система не имеет решений. Плоскости, задаваемые уравнениями

системы, параллельны (рис. 4.9), причем две плоскости (на рис. 4.10 плоскости α_1 и α_2) могут совпадать.

4. Если ранги матриц A и $\overline{A}$ равны между собой и равны двум, то система имеет бесчисленное множество решений. Все три плоскости пересекаются по одной прямой (рис. 4.11). При этом две плоскости могут совпадать (рис. 4.12). Решением системы будут координаты любой точки, принадлежащей прямой.

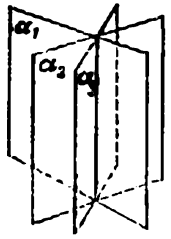


Рис. 4.11

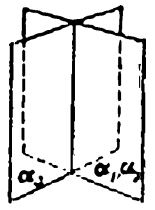


Рис. 4.12

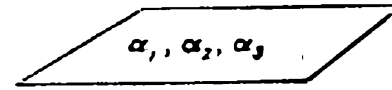


Рис. 4.13

5. Если ранги матриц A и $\overline{A}$ равны между собой и равны единице, то система имеет бесчисленное множество решений. Все три плоскости совпадают (рис. 4.13). Решением системы будут координаты любой точки, принадлежащей совпадающим плоскостям.

§ 9. Неравенства

9.1. Определения и основные свойства неравенств. *Неравенствами* называют выражения вида

$$a < b \quad (a \leqslant b), \qquad a > b \quad (a \geqslant b),$$

где a и b могут быть числами или функциями. Символы $<$ ($\leqslant$), $>$ ($\geqslant$) называются *знаками неравенства* и читаются соответственно: меньше (меньше или равно), больше (больше или равно).

Неравенства, которые записываются с помощью знаков $>$ и $<$, называются *строгими* неравенствами, а неравенства, в записи которых участвуют знаки $\geqslant$ и $\leqslant$, — *нестрогими*.

Нестрогое неравенство эквивалентно строгому неравенству того же знака и равенству.

Различают два вида неравенств: *арифметические* (или числовые), в записи которых участвуют только числа, и *неарифметические*, в записи которых наряду с числами участвуют функции одной или нескольких переменных.

Например, числовыми неравенствами будут

$$2 > 1, \quad \sqrt{2} \leqslant 7.$$

Неарифметическими неравенствами будут, например, неравенства

$$a < 1, \quad \lg^2 x + \operatorname{tg} x > 0, \quad x^2 + y^2 \geqslant R^2.$$

Функции, входящие в неравенства, могут принимать различные числовые значения в зависимости от различных значений своих аргументов. При одних значениях аргументов неравенство может быть справедливым, при других — нет.

Основные свойства неравенств*).

1. если $a > b$, то $b < a$;
если $a < b$, то $b > a$ (антисимметричность).

2. если $a > b$ и $b > c$, то $a > c$;
если $a < b$ и $b < c$, то $a < c$. (транзитивность).

Арифметические действия с неравенствами.

1. Два неравенства одинакового знака можно складывать; при этом в результате сложения получается неравенство того же знака:

если $a > b$ и $c > d$, то $a + c > b + d$;

если $a < b$ и $c < d$, то $a + c < b + d$.

2. Если обе части неравенства умножить (разделить) на одну и ту же положительную величину, то получится неравенство того же знака; если же обе части неравенства умножить (разделить) на отрицательную величину, то получится неравенство противоположного знака:

если $a > b$ и $c > 0$, то $ac > bc$;

если $a < b$ и $c > 0$, то $ac < bc$;

если $a > b$ и $c < 0$, то $ac < bc$;

если $a < b$ и $c < 0$, то $ac > bc$.

3. Как следствие правил сложения неравенств и умножения обеих частей неравенств на одну и ту же величину получается следующее правило вычитания неравенств разного знака.

Из одного неравенства можно почленно (из левой части — левую, а из правой части — правую) вычитать другое неравенство противоположного знака. В результате получится неравенство, имеющее знак первого неравенства:

если $a < b$ и $c > d$, то $a - c < b - d$;

если $a > b$ и $c < d$, то $a - c > b - d$.

4) Если к левой и правой частям неравенства прибавить одну и ту же величину, то в результате получится неравенство того же знака:

если $a > b$, то $a + c > b + c$.

9.2. Некоторые важные неравенства.

1. Для любых действительных чисел a и b выполняется неравенство

$$|a + b| \leqslant |a| + |b|,$$

т. е. абсолютная величина суммы двух действительных чисел не превосходит суммы абсолютных величин этих чисел.

Методом математической индукции можно доказать, что для любого конечного числа слагаемых a_i справедливо неравенство

$$\left|\sum_{i=1}^{n} a_i\right| < \sum_{i=1}^{n} |a_i|,$$

причем равенство достигается в том и только в том случае, когда все слагаемые — числа одного знака.

---

*) Все сказанное ниже для случая строгих неравенств остается справедливым и для нестрогих неравенств.

2. Для любых действительных чисел a и b выполняется неравенство

$$|a-b| \geqslant ||a|-|b||,$$

т. е. абсолютная величина разности двух чисел не меньше абсолютной величины разности абсолютных величин этих чисел.

3. Для любых двух действительных чисел a и b выполняется неравенство

$$a^2+b^2 \geqslant 2|ab|,$$

причем равенство достигается в том и только в том случае, когда $|a|=|b|$.

4. Если a и b — действительные числа одного знака ($ab > 0$), то

$$\frac{a}{b}+\frac{b}{a} \geqslant 2,$$

причем равенство достигается в том и только в том случае, когда $a=b$.

5. *Неравенство Коши*: если a и b — неотрицательные действительные числа, то

$$\frac{a+b}{2} \geqslant \sqrt{ab},$$

т. е. среднее арифметическое двух неотрицательных чисел не меньше их среднего геометрического.

Аналогичное неравенство справедливо и для любого конечного набора неотрицательных чисел a_i ($i=1, 2, 3, \ldots, n$):

$$\frac{a_1+a_2+a_3+\ldots+a_n}{n} \geqslant \sqrt[n]{a_1a_2a_3\ldots a_n},$$

т. е. среднее арифметическое n неотрицательных чисел больше или равно их среднему геометрическому; равенство достигается, когда все n чисел равны.

6. *Неравенство Коши — Буняковского:* для любых действительных чисел $a_1, a_2, a_3, \ldots, a_n, b_1, b_2, \ldots, b_n$ выполняется неравенство

$$(a_1b_1+a_2b_2+a_3b_3+\ldots+a_nb_n)^2 \leqslant \left(a_1^2+a_2^2+\ldots+a_n^2\right)\left(b_1^2+b_2^2+\ldots+b_n^2\right)$$

Равенство достигается тогда и только тогда, когда числа a_i и b_i пропорциональны, т. е. когда существуют такие числа α и β, что $\alpha^2+\beta^2 \neq 0$, и для всех $i=1, 2, 3, \ldots, n$ выполняется равенство

$$\alpha a_i+\beta b_i=0.$$

7. Связь между средним арифметическим и средним квадратичным нескольких чисел: абсолютная величина среднего арифметического любого конечного числа действительных чисел не превосходит среднего квадратичного этих чисел, т. е.

$$\frac{|a_1+a_2+\ldots+a_n|}{n} \leqslant \sqrt{\frac{a_1^2+a_2^2+\ldots+a_n^2}{n}},$$

причем равенство достигается тогда и только тогда, когда все n чисел равны между собой.

8. *Неравенство Гёльдера* (обобщение неравенства Коши—Буняковского): для любых действительных чисел $a_1, a_2, \ldots, a_n, b_1, \ldots, b_n$ при любом $p > 1$

выполняется неравенство

$$|a_1b_1+a_2b_2+\ldots+a_nb_n| \leqslant$$

$$\leqslant (|a_1|^p+|a_2|^p+\ldots+|a_n|^p)^{1/p}(|b_1|^q+|b_2|^q+\ldots+|b_n|^q)^{1/q},$$

где $q=\frac{p}{p-1}$.

При $p=2$ неравенство Гёльдера превращается в неравенство Коши — Буняковского.

§ 10. Решение неравенств и систем неравенств

10.1. Основные определения. Пусть f — числовая функция одной или нескольких переменных (аргументов). Решить неравенство

$$f<0 \tag{1}$$

— это значит найти множество значений аргумента (аргументов) функции f, при которых неравенство (1) справедливо.

Аналогично, решить неравенство

$$f>0 \tag{2}$$

— это значит найти множество значений аргумента (аргументов) функции f, при которых неравенство (2) справедливо.

Множество решений нестрогих неравенств вида $f\leqslant 0$ и $f\geqslant 0$ находится как объединение множеств решений уравнения $f=0$ и соответствующих строгих неравенств.

Множество значений аргументов функции f, при которых соответствующее неравенство справедливо, называют *множеством решений* неравенства или просто *решением* неравенства. Два неравенства считаются *эквивалентными*, если множества их решений совпадают.

Решить систему нескольких неравенств — это значит найти все те значения аргументов функций, входящих в неравенства, при которых все неравенства системы одновременно справедливы.

Две системы неравенств считаются *эквивалентными*, если множества их решений совпадают. Система неравенств s считается эквивалентной двум системам неравенств s_1 и s_2, если множество решений системы s совпадает с объединением множеств решений систем s_1 и s_2.

10.2. Линейные неравенства и системы линейных неравенств с одним неизвестным. Под *линейными* неравенствами понимают неравенства вида

$$ax+b>0,\quad ax+b<0,\quad ax+b\geqslant 0,\quad ax+b\leqslant 0,$$

где a и b — действительные числа ($a\neq 0$).

Линейные неравенства решают заменой исходного неравенства эквивалентным ему. При этом используются следующие преобразования неравенств, приводящие к эквивалентным неравенствам: прибавление к обеим частям неравенства одного и того же числа и умножение (деление) обеих частей неравенства на одно и то же число. Так, множество решений неравенства

$$ax+b>0 \tag{3}$$

может быть найдено следующим образом.

Прибавим к обеим частям неравенства (3) число $-b$, в результате чего получим эквивалентное неравенство

$$ax>-b, \tag{4}$$

и разделим обе части неравенства на a. Тогда:

1) Если $a > 0$, то получаем неравенство

$$x > -\frac{b}{a},$$

которое и дает множество решений исходного неравенства (3). Это множество решений также можно записать в виде

$$x \in \left(-\frac{b}{a};\ +\infty\right).$$

2) Если $a < 0$, то получаем неравенство

$$x < -\frac{b}{a}.$$

Множество действительных чисел, удовлетворяющих этому неравенству, и есть множество решений исходного неравенства (3):

$$x \in \left(-\infty;\ -\frac{b}{a}\right).$$

Аналогичным образом могут быть найдены решения любых линейных неравенств.

Множество решений системы двух линейных неравенств

$$ax + b > 0,$$
$$cx + d > 0$$

находится как пересечение множеств решений этих неравенств.

10.3. Квадратные неравенства. Под *квадратным* неравенством понимается неравенство, которое может быть приведено к одному из следующих неравенств:

$$ax^2 + bx + c > 0, \quad ax^2 + bx + c < 0,$$
$$ax^2 + bx + c \geqslant 0, \quad ax^2 + bx + c \leqslant 0,$$

где a, b, c — некоторые действительные числа и $a \neq 0$.

Простейшими квадратными неравенствами являются неравенства

$$x^2 < m \quad \text{и} \quad x^2 > m.$$

Множество решений неравенства $x^2 < m$:

1) при $m \leqslant 0$ $x = \varnothing$ (т. е. нет решений);

2) при $m > 0$ $x \in (-\sqrt{m};\ \sqrt{m})$, т. е. $-\sqrt{m} < x < \sqrt{m}$, или $|x| < \sqrt{m}$.

Множество решений неравенства $x^2 > m$:

1) при $m < 0$ $x \in \mathbf{R}$ (т. е. x — любое действительное число);

2) при $m \geqslant 0$ $x \in (-\infty;\ -\sqrt{m}) \cup (\sqrt{m};\ +\infty)$, т. е. $-\infty < x < -\sqrt{m}$ и $\sqrt{m} < x < +\infty$, или $|x| > \sqrt{m}$.

Квадратное неравенство $ax^2 + bx + c > 0$ в зависимости от значений своих коэффициентов a, b и c имеет множества решений:

1) при $a > 0$, $D = b^2 - 4ac \geqslant 0$ $x \in \left(-\infty;\ \frac{-b-\sqrt{D}}{2a}\right) \cup$

$\cup \left(\frac{-b+\sqrt{D}}{2a};\ +\infty\right)$;

2) при $a > 0$, $D < 0$ $x \in \mathbf{R}$;

3) при $a < 0$, $D \geqslant 0$ $x \in \left(\frac{-b+\sqrt{D}}{2a}; \frac{-b-\sqrt{D}}{2a}\right)$;

4) при $a < 0$, $D < 0$ $x = \varnothing$ (т. е. решений нет).

Решение неравенства $ax^2+bx+c < 0$ сводится к решению рассмотренного выше неравенства, если обе части неравенства умножить на -1.

Множество решений нестрогих неравенств $ax^2+bx+c \geqslant 0$ и $ax^2+bx+c \leqslant 0$ находится как объединение множеств решений соответствующих строгих неравенств и уравнения $ax^2+bx+c=0$.

Дробно-линейными неравенствами называют неравенства, которые можно привести к виду

$$\frac{ax+b}{cx+d} > k, \tag{5}$$

где a, b, c, d, k — некоторые действительные числа и $c \neq 0$ (если $c=0$, то дробно-линейное неравенство превращается в линейное). К дробно-линейным неравенствам относятся и неравенства вида (5), в которых вместо знака $>$ стоят знаки $<$, $\geqslant$, $\leqslant$.

Решение дробно-линейного неравенства сводится к решению квадратного неравенства. Для этого необходимо умножить обе части неравенства (5) на выражение $(cx+d)^2$, которое положительно при всех $x \in R$ и $x \neq -d/c$.

10.4. Метод интервалов. Пусть $P(x)$ — многочлен n-й степени с действительными коэффициентами $c_1, c_2, \ldots, c_l$ — все действительные корни многочлена с кратностями $k_1, k_2, \ldots, k_l$ соответственно, причем $c_1 > c_2 > \ldots > c_l$. Тогда многочлен $P(x)$ можно представить в виде

$$P(x) = (x-c_1)^{k_1}(x-c_2)^{k_2}\ldots(x-c_l)^{k_l}Q(x), \tag{6}$$

где многочлен $Q(x)$ действительных корней не имеет и либо положителен, либо отрицателен при всех $x \in R$. Положим для определенности, что $Q(x) > 0$. Тогда при $x > c_1$ все сомножители в разложении (6) положительны и $P(x) > 0$. Если c_1 — корень нечетной кратности (k_1 нечетное), то при $c_2 < x < c_1$ все сомножители в разложении (6) за исключением первого, положительны и $P(x) < 0$. В этом случае говорят, что многочлен $P(x)$ меняет свой знак при переходе через корень c_1. Если же c_1 — корень четной кратности (k_1 четное), то все сомножители (в том числе и первый при $c_2 < x < c_1$) положительны и, следовательно, $P(x) > 0$. В этом случае говорят, что многочлен $P(x)$ не меняет своего знака при переходе через корень c_1.

Аналогичным способом, используя разложение (6), нетрудно убедиться, что при переходе через корень c_2 многочлен $P(x)$ меняет знак, если k_2 нечетно, и не меняет знака, если k_2 четно.

Рассмотренное свойство многочленов используется для решения неравенств методом интервалов. Для того чтобы найти все решения неравенства

$$P(x) > 0,$$

достаточно знать все действительные корни многочлена $P(x)$, их кратности и знак многочлена $P(x)$ в произвольно выбранной точке x_0, не совпадающей с корнем многочлена.

Пример 1. Решить неравенство

$$x^2(x+2)(x-1)^3(x^2+1) > 0. \tag{7}$$

Отметим на числовой оси Ox корни многочлена, стоящего в левой части неравенства (рис. 4.14).

При $x > 1$ многочлен положителен, так как все сомножители, стоящие в левой части равенства, положительны. Будем двигаться по оси Ox справа налево. При переходе через точку $x = 1$ многочлен меняет свой знак и становится отрицательным, так как $x = 1$ — корень кратности 3; при переходе через точку $x = 0$ многочлен знака не меняет, так как $x = 0$ — корень кратности 2; при переходе через точку $x = -2$ многочлен опять меняет знак и становится положительным.

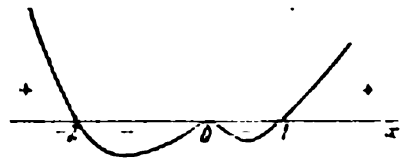


Рис 4.14

Промежутки знакопостоянства данного многочлена схематически изображены на рис. 4.14. Используя этот рисунок, легко выписать множество решений неравенства (7):

$$x \in (-\infty; -2) \cup (1; +\infty).$$

Рациональное неравенство вида

$$\frac{P(x)}{Q(x)} > 0, \tag{8}$$

где $P(x)$ и $Q(x)$ — многочлены, может быть также решено методом интервалов. Умножая обе части неравенства (8) на многочлен $[Q(x)]^2$, который положителен при всех допустимых значениях неравенства (8), получаем неравенство

$$P(x)\,Q(x) > 0,$$

эквивалентное неравенству (8).

Пример 2. Решить рациональное неравенство

$$\frac{x^3 (x-1)^2 (x+2)}{(x-3)^3} < 0. \tag{9}$$

Умножив обе части неравенства на $(x-3)^4$, получим неравенство, эквивалентное неравенству (9):

$$x^3 (x-1)^2 (x+2)(x-3) < 0.$$

Множество решений последнего неравенства находится методом интервалов и будет следующим:

$$x \in (-\infty; -2) \cup (1; 3).$$

10.5. Решение иррациональных неравенств. Под *иррациональным* неравенством понимается неравенство, в котором неизвестные величины (или рациональные функции неизвестных величин) находятся под знаком радикала. Для того чтобы найти множество решений иррационального неравенства, приходится, как правило, возводить обе части неравенства в натуральную степень. Несмотря на внешнюю схожесть процедур решения иррационального уравнения (см. п. 5.9) и иррационального неравенства, между ними существует большое отличие. При решении иррациональных уравнений можно не заботиться о том, что после возведения в степень получилось уравнение, эквивалентное исходному: алгебраическое уравнение имеет конечное число корней, из которых проверкой нетрудно отобрать решения исходного иррационального уравнения.

Множество решений неравенства представляет собой, как правило, бесконечное множество чисел, и поэтому непосредственная проверка решений путем подстановки этих чисел в исходное неравенство ста-

новится принципиально невозможной. Единственный способ, гарантирующий правильность ответа, заключается в том, что мы должны следить за тем, чтобы при каждом преобразовании неравенства у нас получалось неравенство, эквивалентное исходному.

Решая иррациональные неравенства, следует помнить, что при возведении обеих его частей в нечетную степень всегда получается неравенство, эквивалентное исходному неравенству. Если же обе части неравенства возводить в четную степень, то будет получаться неравенство, эквивалентное исходному и имеющее тот же знак, лишь в случае, если обе части исходного неравенства неотрицательны.

Пример. Решить неравенство

$$\sqrt{x-5}-\sqrt{9-x}>1. \tag{10}$$

Множеством допустимых значений x, которое находится как решение системы

$$x-5\geqslant 0,$$
$$9-x\geqslant 0,$$

является промежуток $[5;\ 9]$.

Неравенство (10) эквивалентно неравенству

$$\sqrt{x-5}>1+\sqrt{9-x}, \tag{11}$$

обе части которого неотрицательны. Возводя обе части неравенства (11) в квадрат, в результате несложных преобразований получаем неравенство

$$2x-15>2\sqrt{9-x}, \tag{12}$$

эквивалентное неравенству (11). Далее возможны два случая:

1) если $2x-15\leqslant 0$, т. е. $x\leqslant 15/2$, левая часть неравенства отрицательна или равна нулю, а правая неотрицательна. Поэтому ни при каком $x\in[5;\ 15/2]$ неравенство (12) не выполняется;

2) если $2x-15>0$, то обе части неравенства неотрицательны и после возведения в квадрат получаем неравенство, эквивалентное неравенству (12):

$$(2x-15)^2>4(9-x).$$

Таким образом, множество решений неравенства (10) получается как множество решений системы неравенств

$$5\leqslant x\leqslant 9,$$
$$2x-15>0,$$
$$(2x-15)^2>4(9-x),$$

откуда получаем $x\in\left(\dfrac{14+\sqrt{7}}{2};\ 9\right]$.

10.6. Показательные неравенства. Простейшим *показательным* неравенством является неравенство вида

$$a^x>b, \tag{13}$$

где a и b — некоторые действительные числа ($a>0$, $a\neq 1$).

В зависимости от значений параметров a и b множество решений неравенства (13) будет следующим:

1) при $a>1$, $b>0$ $x\in(\log_a b;\ +\infty)$;

2) при $0<a<1$, $b>0$ $x\in(-\infty;\ \log_a b)$;

3) при $a>0$, $b<0$ $x\in\mathbb{R}$.

Множество решений неравенства

$$a^x < b \tag{14}$$

в зависимости от значений параметров a и b будет следующим:

1) при $a > 1$, $b > 0$ $x \in (-\infty; \log_a b)$;

2) при $0 < a < 1$, $b > 0$ $x \in (\log_a b; +\infty)$;

3) при $a > 0$, $b \leqslant 0$ $x \in \varnothing$ (т. е. неравенство решений не имеет).

Неравенства вида (13) и (14) могут быть обобщены на случай, когда в показателе степени стоит некоторая функция от x. Так, например, множество решений неравенства

$$2^{f(x)} > 3 \tag{15}$$

находится как множество решений неравенства

$$f(x) > \log_2 3,$$

эквивалентного неравенству (15).

Методы сведения более сложных показательных неравенств к неравенствам вида (13), (14) аналогичны методам, используемым при решении показательных уравнений (см. п. 6.1). Так, например, решение показательного неравенства вида $P(a^x) > 0$, где $P(x)$ — многочлен указанного аргумента, заменой $a^x = y$ сводится к последовательному решению неравенства $P(y) > 0$ и решению простейших показательных неравенств вида (13), (14) или систем простейших показательных неравенств.

Пример. Решить неравенство

$$9^x - 10 \cdot 3^x + 9 \leqslant 0.$$

Обозначим $3^x = y$. Так как $9^x = (3^2)^x = (3^x)^2$, то $9^x = y^2$ и данное неравенство для неизвестной y принимает вид $y^2 - 10y + 9 \leqslant 0$. Решение этого квадратного неравенства: $1 \leqslant y \leqslant 9$. Это двойное неравенство эквивалентно системе двух неравенств

$$\begin{cases} y \geqslant 1, \\ y \leqslant 9, \end{cases}$$

которая для неизвестной x имеет вид

$$\begin{cases} 3^x \geqslant 1, \\ 3^x \leqslant 9 \end{cases} \Leftrightarrow \begin{cases} x \geqslant 0, \\ x \leqslant 2 \end{cases} \quad \text{или} \quad 0 \leqslant x \leqslant 2.$$

10.7. Логарифмические неравенства. Простейшими *логарифмическими* неравенствами являются неравенства вида

$$\log_a x > b, \tag{16}$$

$$\log_a x < b, \tag{17}$$

где a и b — некоторые действительные числа ($a > 0$, $a \neq 1$).

В зависимости от значений a множества решений неравенства (16) будут следующими:

1) при $a > 1$ $x \in (a^b; +\infty)$;

2) при $0 < a < 1$ $x \in (0; a^b)$,

а неравенства (17) — следующими:

1) при $a > 1$ $x \in (0; a^b)$;

2) при $0 < a < 1$ $x \in (a^b; +\infty)$.

Неравенства вида (16), (17) могут быть обобщены на случай, когда аргументом логарифмической функции является некоторая функция

$f(x)$. Так, например, логарифмическое неравенство вида $\log_a f(x) > b$ эквивалентно следующим системам неравенств *):

1) при $a > 1$ $f(x) > 0$, $f(x) > a^b$;
2) при $0 < a < 1$ $f(x) > 0$, $f(x) < a^b$.

Неравенство вида $\log_a f(x) < b$ эквивалентно следующим системам неравенств:

1) при $a > 1$ $f(x) > 0$, $f(x) < a^b$;
2) при $0 < a < 1$ $f(x) > 0$, $f(x) > a^b$.

Более сложные логарифмические неравенства сводятся к неравенствам вида (16), (17) методами, аналогичными используемым при решении логарифмических уравнений (см. п. 6.2).

Так, например, множество решений неравенства вида

$$P(\log_a x) > 0, \tag{18}$$

а также неравенств $P < 0$, $P \geqslant 0$, $P \leqslant 0$, где P — многочлен указанного аргумента, находится следующим образом: вводится новая неизвестная $y = \log_a x$ и неравенство (18) решается как алгебраическое относительно неизвестной y. После этого решение исходного неравенства сводится к решению соответствующих простейших неравенств (16), (17) или систем этих неравенств.

Пример. Решить логарифмическое неравенство

$$\log_4^2 x - \log_4 x - 15 > 0. \tag{19}$$

Множество допустимых значений неизвестной данного неравенства: $x > 0$. Приведя все логарифмы к одному основанию (скажем, к основанию 2), получаем неравенство, эквивалентное неравенству (19):

$$\frac{1}{4}\log_2^2 x - \log_2 x - 15 > 0. \tag{20}$$

Решаем это неравенство как квадратное относительно новой неизвестной $y = \log_2 x$. В результате получаем, что логарифмическое неравенство (20) эквивалентно двум простейшим неравенствам: $\log_2 x > 10$ и $\log_2 x < -6$. Объединение их множеств решений и дает множество решений неравенства (19):

$$x \in (0;\ 2^{-6}) \cup (2^{10};\ +\infty).$$

В заключение отметим еще один часто встречающийся тип логарифмических неравенств, в которых как основание логарифма, так и аргумент логарифма являются функциями неизвестной x. Таким логарифмическим неравенством будет, например, неравенство

$$\log_{g(x)} f(x) > c, \tag{21}$$

где $f(x)$ и $g(x)$ — некоторые многочлены, а c — некоторое действительное число. Множество допустимых значений неизвестной x находится как множество решений системы

$$f(x) > 0,$$
$$g(x) > 0,$$
$$g(x) \neq 1.$$

*) В случае, если исходное неравенство нестрогое, вторые неравенства этих систем также нестрогие.

Логарифмическое неравенство (21) эквивалентно двум системам алгебраических неравенств:

$$\begin{array}{ll} f(x) > 0, & f(x) > 0, \\ g(x) > 1, & 0 < g(x) < 1, \\ f(x) > [g(x)]^c; & f(x) < [g(x)]^c. \end{array}$$

Пример. Решить неравенство

$$\log_x \left(2x - \frac{3}{4}\right) > 2. \tag{22}$$

Данное логарифмическое неравенство эквивалентно двум системам неравенств:

$$\begin{array}{ll} 2x - \frac{3}{4} > 0, & 2x - \frac{3}{4} > 0, \\ x > 1, & 0 < x < 1, \\ 2x - \frac{3}{4} > x^2; & 2x - \frac{3}{4} < x^2. \end{array}$$

Множество решений логарифмического неравенства (22) находим как объединение множеств решений этих двух систем и окончательно получаем

$$x \in \left(\frac{1}{5};\ \frac{1}{2}\right) \cup \left(1;\ \frac{3}{2}\right).$$

10.8. Геометрическое изображение множества решений неравенства с двумя неизвестными. Пусть имеется некоторое неравенство с двумя неизвестными. Множество решений такого неравенства — это множество упорядоченных пар действительных чисел $(x; y)$, которое геометрически может быть изображено как множество точек координатной плоскости Oxy.

Объединение множеств решений неравенств

$$F(x, y) > 0, \qquad F(x, y) < 0$$

и уравнения

$$F(x, y) = 0$$

дает множество всех точек координатной плоскости Oxy, за исключением точек, в которых функция двух переменных $F(x, y)$ не определена. Так, например, объединение множеств решений неравенств

$$\frac{x}{y} > 0, \qquad \frac{x}{y} < 0$$

и уравнения

$$\frac{x}{y} = 0$$

дает множество всех точек числовой плоскости $\mathbf{R}^2$, за исключением множества точек, принадлежащих оси Ox.

Уравнение $F(x, y) = 0$ задает множество точек «границы раздела» (либо часть этой границы) между множествами точек, задаваемых неравенствами $F(x, y) > 0$ и $F(x, y) < 0$.

Если уравнение $F(x, y) = 0$ имеет единственное решение относительно переменной y (x считается параметром), то говорят, что уравнение $F(x, y) = 0$ задает функцию

$$y = f(x).$$

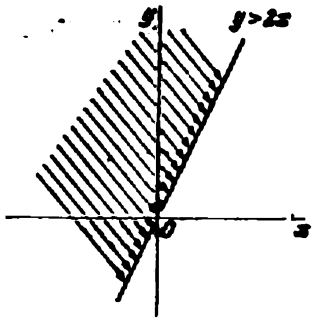


Рис. 4.15

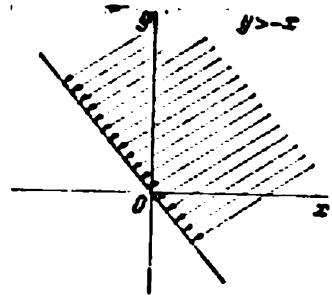


Рис. 4.16

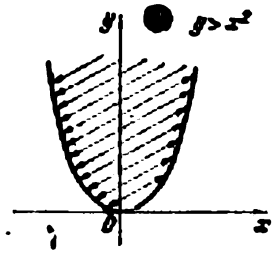


Рис. 4.17

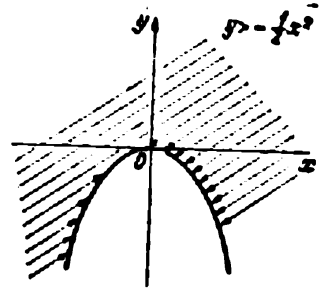


Рис. 4.18

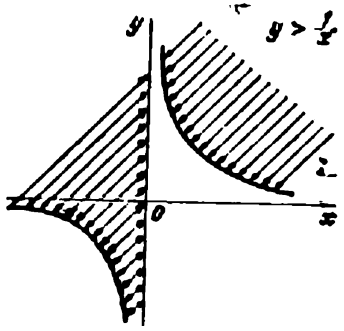


Рис. 4.19

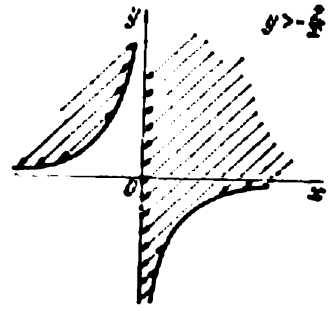


Рис. 4.20

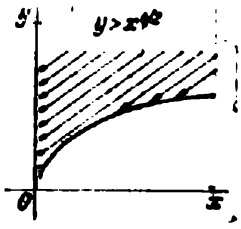


Рис. 4.21

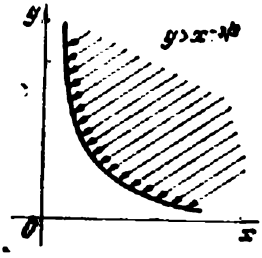


Рис. 4.22

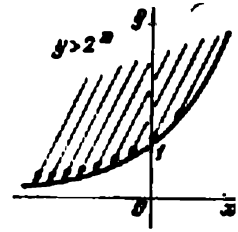


Рис. 4.23

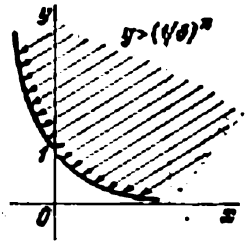


Рис. 4.24

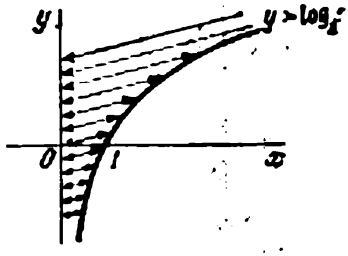


Рис. 4.25

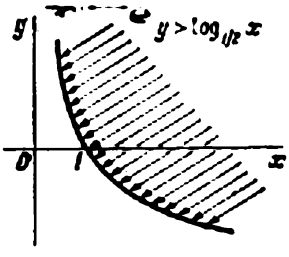


Рис. 4.26

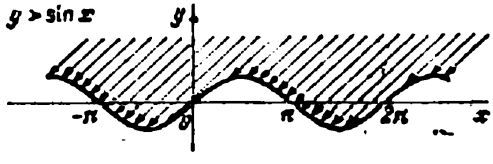


Рис. 4.27

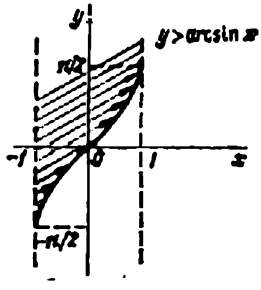


Рис. 4.28

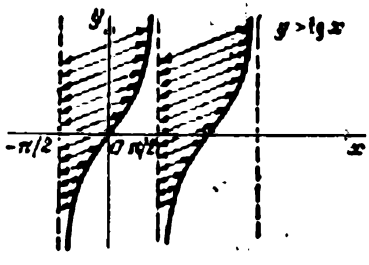


Рис. 4.29

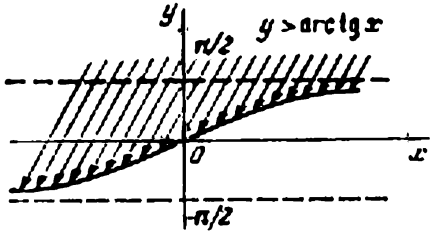


Рис. 4.30

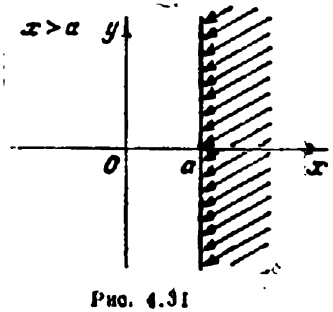


Рис. 4.31

Если уравнение $F(x, y)=0$ имеет единственное решение относительно переменной x (y считается параметром), то говорят, что уравнение $F(x, y)=0$ задает функцию

$$x=g(y),$$

где y считается независимой переменной, а x — зависимой.

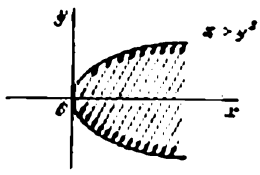


Рис. 4.32

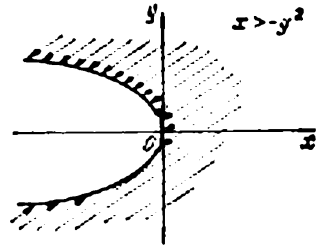


Рис. 4.33

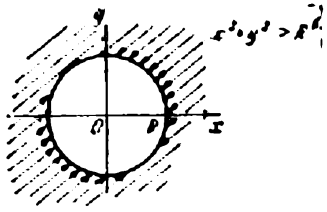


Рис. 4.34

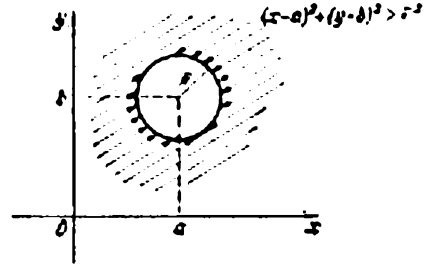


Рис. 4.35

Может оказаться, что уравнение $F(x, y)=0$ не имеет единственного решения ни относительно переменной x, ни относительно переменной y. В этом случае говорят, что уравнение $F(x, y)=0$ задает соответствие между множествами допустимых значений переменных x и y.

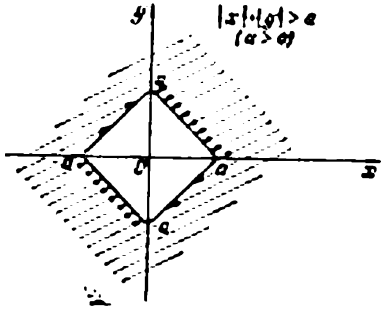


Рис. 4.36

На рис. 4.15—4.30 изображены некоторые фигуры, задаваемые неравенствами вида $y > f(x)$. Эти фигуры представляют собой часть координатной плоскости Oxy, лежащую выше линии $y=f(x)$. Фигуры, задаваемые неравенствами $y < f(x)$, представляют собой часть плоскости, лежащую ниже линии $y=f(x)$.

На рис. 4.31—4.33 заштрихованы фигуры, задаваемые неравенствами вида $x > g(y)$.

На рис. 4.34—4.36 изображены фигуры, задаваемые неравенствами вида $F(x, y) > 0$, в случае, когда уравнение $F(x, y)=0$ задает соответствие между переменными x и y:

1) неравенство $x^2+y^2 > R^2$ задает внешность круга радиуса R с центром в начале координат (см. рис. 4.34);

2) неравенство $(x-a)^2+(y-b)^2 > R^2$ задает внешность круга радиуса R с центром в точке с координатами $(a; b)$ (см. рис. 4.35);

3) неравенство вида $|x|+|y|>a$ $(a>0)$ задает внешность квадрата с вершинами в точках $(a; 0)$, $(0; a)$, $(-a; 0)$ и $(0; -a)$ (см. рис. 4.36).

На рис. 4.15—4.36 изображены простейшие фигуры, задаваемые неравенствами с двумя переменными x и y. Однако умение строить эти простейшие фигуры позволяет строить и более сложные фигуры, задаваемые, например, системой нескольких неравенств с двумя переменными. Геометрическая фигура, задаваемая системой неравенств с двумя переменными, представляет собой общую часть (пересечение) всех фигур, задаваемых неравенствами системы.

§ 11. Приемы доказательства справедливости неравенств

11.1. Доказательство справедливости неравенств с помощью цепочки эквивалентных неравенств. Один из методов доказательства справедливости неравенств основан на построении цепочки эквивалентных неравенств, в конце которой получается неравенство, справедливость которого очевидна.

Пример 1. Доказать неравенство

$$\frac{a^3+b^3}{2}\geqslant\left(\frac{a+b}{2}\right)^3, \text{ где } a>0 \text{ и } b>0.$$

Заменим это неравенство на эквивалентное неравенство

$$\frac{a^3+b^3}{2}-\left(\frac{a+b}{2}\right)^3\geqslant 0.$$

Раскрывая скобки и группируя слагаемые левой части неравенства, получим эквивалентное неравенство

$$\frac{3}{8}(a+b)(a-b)^2\geqslant 0.$$

Так как $a>0$ и $b>0$, то справедливость последнего неравенства очевидна, что и доказывает справедливость исходного неравенства.

Пример 2. Доказать неравенство

$$\frac{1}{\log_2\pi}+\frac{1}{\log_\pi 2}>2.$$

На основании свойства логарифмов

$$\frac{1}{\log_\pi 2}=\log_2\pi$$

исходное неравенство преобразуется к эквивалентному неравенству

$$\log_2\pi+\frac{1}{\log_2\pi}>2. \qquad (1)$$

Так как $\log_2\pi>0$, то в левой части неравенства (1) стоит сумма двух взаимно обратных положительных чисел, отличных от единицы. В силу неравенства 4) п. 9.2 неравенство (1) справедливо, а следовательно, справедливо и исходное неравенство.

Пример 3. Доказать, что для любых действительных x и y выполняется неравенство

$$x^2+2xy+3y^2+2x+6y+3\geqslant 0.$$

Преобразуем левую часть данного неравенства:

$$x^2+2xy+3y^2+2x+6y+3=$$
$$=(x^2+2xy+y^2)+2y^2+2x+6y+3=$$
$$=(x+y)^2+2(x+y)+2y^2+4y+3=$$
$$=[(x+y)^2+2(x+y)+1]+2y^2+4y+2=$$
$$=[(x+y)+1]^2+2(y^2+2y+1)=[(x+y)+1]^2+2(y+1)^2.$$

В результате проведенных тождественных преобразований исходное неравенство приводится к эквивалентному неравенству

$$(x+y+1)^2+2(y+1)^2\geqslant 0,$$

справедливость которого очевидна.

Аналогичным образом, используя неравенства п. 9.2, можно доказать, например, следующие неравенства:

1) $a^2+b^2+c^2+3\geqslant 2(a+b+c)$;

2) $\frac{bc}{a}+\frac{ac}{b}+\frac{ab}{c}\geqslant a+b+c \quad (a>0,\ b>0,\ c>0)$;

3) $(a+b)(b+c)(a+c)\geqslant 8abc \quad (a>0,\ b>0,\ c>0)$;

4) $a^2+b^2+c^2\geqslant ab+bc+ac$ (a, b, c — числа одного знака).

11.2. Доказательство справедливости неравенств с использованием свойств функций, входящих в неравенства. Справедливость некоторых неравенств может быть доказана с использованием свойств функций, входящих в неравенства.

Так, например, справедливость неравенства

$$\operatorname{tg} x+\operatorname{ctg} x\geqslant 2$$

для всех $x\in(0;\ \pi/2)$ следует из формулы, связывающей тригонометрические функции $\operatorname{tg} x$ и $\operatorname{ctg} x$:

$$\operatorname{ctg} x=\frac{1}{\operatorname{tg} x}.$$

Действительно, в силу этой формулы исходное неравенство запишется в виде неравенства

$$\operatorname{tg} x+\frac{1}{\operatorname{tg} x}\geqslant 2,$$

которое справедливо для всех значений $x\in(0;\ \pi/2)$, так как в этом промежутке $\operatorname{tg} x$ положителен, а сумма двух любых положительных взаимно обратных чисел больше или равна двум.

Другим примером неравенства, справедливость которого доказывается на основании свойств входящих в него функций, является неравенство

$$0\leqslant \sin^8 x+\cos^{20} x\leqslant 1.$$

Справедливость этого неравенства, а также, например, неравенства

$$\sqrt[4]{\sin x}+\sqrt[4]{\cos x}\geqslant 1$$

следует из ограниченности функций $y=\sin x$, $y=\cos x$ и соотношений п. 3.2 гл. 6.

В следующем примере доказательство справедливости неравенства основано на использовании свойства монотонности показательной функции.

Докажем, что если $a+b=c$ $(a>0,\ b>0)$, то:

1) $a^\alpha+b^\alpha>c^\alpha$, где $\alpha<1$;

2) $a^\beta+b^\beta<c^\beta$, где $\beta>1$.

По условию задачи

$$\left\{\begin{array}{l}\dfrac{a}{c}+\dfrac{b}{c}=1 \quad \text{и} \quad \dfrac{a}{c}>0, \quad \dfrac{b}{c}>0.\end{array}\right.$$

Если сумма двух положительных чисел равна единице, то каждое из этих чисел меньше единицы:

$$0<\frac{a}{c}<1, \quad 0<\frac{b}{c}<1.$$

В силу свойства монотонного убывания показательной функции с основанием, меньшим единицы, имеем

$$\begin{array}{l}\left(\dfrac{a}{c}\right)^\alpha>\dfrac{a}{c}, \quad \left(\dfrac{b}{c}\right)^\alpha>\dfrac{b}{c},\\ \left(\dfrac{a}{c}\right)^\beta<\dfrac{a}{c}, \quad \left(\dfrac{b}{c}\right)^\beta<\dfrac{b}{c},\end{array} \tag{2}$$

где α — любое действительное число, меньшее единицы, а β — действительное число, большее единицы. Складывая попарно неравенства (2), получаем

$$\left(\frac{a}{c}\right)^\alpha+\left(\frac{b}{c}\right)^\alpha>\frac{a}{c}+\frac{b}{c}=1\Leftrightarrow a^\alpha+b^\alpha>c^\alpha,$$

$$\left(\frac{a}{c}\right)^\beta+\left(\frac{b}{c}\right)^\beta<\frac{a}{c}+\frac{b}{c}=1\Leftrightarrow a^\beta+b^\beta<c^\beta,$$

что и требовалось доказать.

11.3. Некоторые специальные приемы доказательства справедливости неравенств.

1. Один из таких приемов состоит в следующем. Пусть, например, требуется доказать справедливость неравенства $A<B$. Если удается подобрать такую величину C, что $A<C$, и доказать справедливость неравенства $C\leqslant B$, то тем самым будет доказана справедливость исходного неравенства $A<B$. Именно такой прием использовался при доказательстве ограниченности последовательности (см. п. 2.4 гл. 8)

$$x_n=\left(1+\frac{1}{n}\right)^n.$$

2. Справедливость некоторых неравенств, в которых левая и правая части — функции натурального аргумента, может быть доказана следующим приемом. Исходное неравенство преобразуется к эквивалентному неравенству, у которого правая часть постоянная, а левая есть функция натурального аргумента. Тогда, рассматривая функцию натурального аргумента, стоящую в левой части неравенства, как общий член некоторой последовательности, доказательство справедливости неравенства можно свести к исследованию свойств этой последовательности.

Пример. Доказать, что неравенство

$$2^{n-1}\leqslant n!$$

справедливо при любом натуральном n.

Преобразуем данное неравенство к эквивалентному неравенству

$$\frac{2^{n-1}}{n!} \leqslant 1. \qquad (3)$$

Рассмотрим последовательность (x_n), задаваемую формулой общего члена

$$x_n = \frac{2^{n-1}}{n!}.$$

Нетрудно доказать, что эта последовательность — монотонно убывающая и, следовательно, наибольший член данной последовательности — это первый член, который равен единице. Но так как при $n = 1$ неравенство (3) верно, то оно будет верно и при всех остальных значениях n.

3. Справедливость некоторых неравенств может быть доказана методом математической индукции. Именно этот метод использован для доказательства ограниченности последовательности (см. пример 3 п. 2.4 гл. 8)

$$x_n = \underbrace{\sqrt{c+\sqrt{c+\ldots+\sqrt{c}}}}_{n \text{ корней}}.$$

Методом математической индукции могут быть также доказаны, например, следующие неравенства:

1) $n! > \left(\frac{n+1}{n}\right)^n$, если $n > 2$;

2) $|\sin nx| \leqslant n|\sin x|$;

3) $n! > 2^{n-1}$, если $n > 2$;

4) $2^n n! < n^n$, если $n > 2$.

11.4. Некоторые способы проверки справедливости числовых неравенств. При решении некоторых уравнений и неравенств иногда необходимо выяснить, какое из двух (или нескольких) чисел наибольшее, а какое — наименьшее. Ниже на примерах рассмотрены некоторые простейшие способы проверки справедливости числовых неравенств.

1. Для того чтобы выяснить, какое из двух рациональных чисел m/n и p/q больше, следует вычислить разность этих чисел:

$$\frac{m}{n} - \frac{p}{q} = \frac{mq - np}{nq}.$$

Если разность больше нуля, то $\frac{m}{n} > \frac{p}{q}$; если разность меньше нуля, то $\frac{m}{n} < \frac{p}{q}$.

2. При сравнении двух иррациональных чисел a и b часто используют следующий способ: предполагают, что одно число больше другого (скажем, $a > b$), и строят цепочку эквивалентных неравенств, приводящую либо к неравенству, справедливость которого очевидна (и тогда предположение, что $a > b$, верно), либо к неравенству, которое заведомо несправедливо (в этом случае сделанное предположение, что $a > b$, неверно, и, следовательно, $a < b$). Например, пусть требуется выяснить, какое число больше: $\sqrt{14} - \sqrt{3}$ или 2. Пред-

положим, что $\sqrt{14}-\sqrt{3}$ меньше 2:

$$\sqrt{14}-\sqrt{3}<2. \tag{4}$$

Так как обе части неравенства (4) положительны, то их можно возвести в квадрат и сохранить тот же знак, который был у предыдущего неравенства: $(\sqrt{14}-\sqrt{3})^2<4$. Раскрывая скобки, получаем $14-2\sqrt{42}+3<4$. Последнее неравенство можно переписать в виде $13<2\sqrt{42}$. Возведя обе части неравенства в квадрат, получим неравенство $169<168$. Очевидно, что последнее неравенство несправедливо, следовательно, предположение о том, что число $\sqrt{14}-\sqrt{3}$ меньше 2, неверно. Число $\sqrt{14}-\sqrt{3}$ больше 2.

3. Разберем на двух примерах, как можно проверить справедливость неравенства с помощью приближенных вычислений.

Пример 1. Доказать, что

$$1/\sqrt{3}<\sin 39^\circ. \tag{5}$$

Попытаемся найти такое число a, чтобы $a<\sin 39^\circ$ и $1/\sqrt{3}<a$. Ближайшее «табличное» значение функции $\sin x$ меньшее, чем $\sin 39^\circ$,— это $\sin 30^\circ=1/2$:

$$\sin 30^\circ<\sin 39^\circ. \tag{6}$$

Однако, выбирая a равным $1/2$, мы ничего не докажем, так как $1/2<1/\sqrt{3}$. Попытаемся улучшить оценку $\sin 39^\circ$. Это можно сделать двумя способами. Первый способ заключается в том, что, последовательно вычисляя с помощью формул половинного угла $\sin 15^\circ$, $\cos 15^\circ$, $\sin 7^\circ30'$, $\cos 7^\circ30'$, мы можем вычислить

$$\sin 37^\circ30'=\sin 30^\circ\cos 7^\circ30'+\sin 7^\circ30'\cos 30^\circ$$

и, учитывая неравенство

$$\sin 37^\circ30'<\sin 39^\circ, \tag{7}$$

взяв в качестве a число $\sin 37^\circ30'$, доказать, что $37^\circ30'>1/\sqrt{3}$.

Однако, как будет показано ниже, нам достаточно иметь более грубую оценку, нежели оценка (7). В п. 1.6 гл. 6 вычислен $\sin 36^\circ$:

$$\sin 36^\circ=\frac{\sqrt{10-2\sqrt{5}}}{4}.$$

Так как в первой четверти $\sin x$ — монотонно возрастающая функция, то из неравенства $36^\circ<39^\circ$ следует неравенство $\sin 36^\circ<\sin 39^\circ$, т. е.

$$\sin 39^\circ>\frac{\sqrt{10-2\sqrt{5}}}{4}.$$

Способом, изложенным выше в п. 2, докажем, что

$$\sin 36^\circ=\frac{\sqrt{10-2\sqrt{5}}}{4}>\frac{1}{\sqrt{3}}.$$

Справедливость последнего неравенства следует из приведенных ниже вычислений:

$$\left(\frac{\sqrt{10-2\sqrt{5}}}{4}\right)^2 > \frac{1}{3}, \quad \frac{5-\sqrt{5}}{8} > \frac{1}{3},$$

$$15-3\sqrt{5} > 8, \quad 7 > 3\sqrt{5}, \quad 49 > 45.$$

Таким образом, доказано, что если положить a равным $\sin 36°$, то

$$\sin 39° > \sin 36°, \quad \text{и} \quad \sin 36° > 1/\sqrt{3}.$$

Из этих двух неравенств следует неравенство $1/\sqrt{3} < \sin 39°$, справедливость которого и требовалось доказать.

Пример 2. Сравнить, какое из двух чисел больше: $\log_2 3$ или $\log_3 5$.

Числа 3 и 5, логарифмы которых требуется сравнить, заключены между следующими степенями чисел 2 и 3, являющихся соответственно основаниями первого и второго логарифмов:

$$2^1 < 3 < 2^2, \quad 3^1 < 5 < 3^2. \tag{8}$$

Так как функция $\log_a x$ при $a > 1$ — монотонно возрастающая функция, из неравенств (8) следуют неравенства

$$\log_2(2^1) < \log_2 3 < \log_2(2^2), \quad \log_3(3^1) < \log_3 5 < \log_3(3^2),$$
$$1 < \log_2 3 < 2, \qquad 1 < \log_3 5 < 2;$$

однако на основании этих неравенств нельзя сделать вывод, какой из логарифмов больше, так как они «зажаты» между одними и теми же числами.

Улучшим оценки (8), оценивая числа 3 и 5 дробными степенями чисел 2 и 3 соответственно. Сравним с числами $2^{1+1/2} = 2\sqrt{2}$ и $3^{1+1/2} = 3\sqrt{3}$ числа 3 и 5 соответственно. С помощью способа, изложенного в п. 2, можно показать, что

$$3 > 2\sqrt{2} \quad \text{и} \quad 5 < 3\sqrt{3}.$$

Таким образом, имеем

$$2^{3/2} < 3 < 2^2, \quad 3^1 < 5 < 3^{3/2},$$

откуда соответственно следует, что

$$\frac{3}{2} < \log_2 3 < 2, \quad 1 < \log_3 5 < \frac{3}{2}. \tag{9}$$

Сравнивая последние неравенства, видим что $\log_2 3$ больше $\log_3 5$.

Заметим, что в случае необходимости оценки (9) можно улучшать и далее. Так, доказывая справедливость неравенств

$$2^{\frac{3}{2}+\frac{1}{4}} = 2\cdot\sqrt{2}\cdot\sqrt[4]{2} > 3, \quad 3^{1+\frac{1}{4}} = 3\sqrt[4]{3} < 5,$$

получаем следующие оценки для $\log_2 3$ и $\log_3 5$:

$$\frac{3}{2} < \log_2 3 < \frac{7}{4}, \quad \frac{5}{4} < \log_3 5 < \frac{3}{2}.$$

Глава 5

ГЕОМЕТРИЯ

Возникновение геометрии было обусловлено практическими потребностями человеческой деятельности (необходимостью измерения земельных участков, измерения объемов различных тел и т. д.). Простейшие геометрические сведения и понятия были известны еще в Древнем Египте. В этот период геометрические утверждения формулировались в виде правил, даваемых без доказательств. С VII века до н. э. по I век н. э. геометрия как наука бурно развивалась в Древней Греции. В этот период происходило не только накопление различных геометрических сведений, но и отрабатывались методика доказательств геометрических утверждений, а также делались первые попытки сформулировать основные первичные положения (аксиомы) геометрии, из которых чисто логическими рассуждениями выводится множество различных геометрических утверждений.

Уровень развития геометрии в Древней Греции отражен в сочинении Евклида «Начала». В этой книге впервые была сделана попытка дать систематическое построение планиметрии на базе основных неопределяемых геометрических понятий и аксиом (постулатов). Особое место в истории математики занимает пятый постулат Евклида (аксиома о параллельных прямых). Долгое время математики безуспешно пытались вывести пятый постулат из остальных постулатов Евклида и лишь в середине XIX века благодаря исследованиям Н. И. Лобачевского, Б. Римана и Я. Бойяи стало ясно, что пятый постулат не может быть выведен из остальных, а система аксиом, предложенная Евклидом,— не единственно возможная. «Начала» Евклида оказали огромное влияние на развитие математики. Эта книга на протяжении более чем двух тысяч лет была не только учебником по геометрии, но и служила отправным пунктом для очень многих математических исследований, в результате которых возникли новые самостоятельные разделы математики.

Систематическое построение геометрии обычно производится по следующему плану.

I. Перечисляются основные геометрические понятия, которые вводятся без определений.

II. Дается формулировка аксиом геометрии.

III. На основе аксиом и основных геометрических понятий формулируются остальные геометрические понятия и теоремы.

§ 1. Луч. Отрезок

1.1. Луч. Пусть a — некоторая прямая, O — некоторая точка прямой a. Точка O разбивает множество точек прямой на два множества: множество точек, лежащих левее точки O, и множество точек, лежащих правее точки O (рис. 5.1). Эти множества называют *открытыми лучами*, исходящими из точки O (или открытыми лучами с

началом в точке O). Множество всех точек прямой a, лежащих правее (левее) точки O, включая точку O, называют *лучом* и обозначают Oa (указывая прямую и начало луча) *). О луче Oa говорят, что он принадлежит прямой a.

Возможны следующие случаи взаимного расположения двух лучей O_1h_1 и O_2h_2.

1. Лучи O_1h_1 и O_2h_2 принадлежат одной прямой. Говорят, что лучи O_1h_1 и O_2h_2 *сонаправлены*, если один из лучей содержится в другом (рис. 5.2), т. е. их пересечением является луч, и пишут

Рис. 5.1

$$O_1h_1 \uparrow\uparrow O_2h_2.$$

Говорят, что лучи O_1h_1 и O_2h_2 *противоположно направлены*, если ни один из них не содержится в другом (рис. 5.3), т. е. их пересечение не является лучом, и пишут

$$O_1h_1 \uparrow\downarrow O_2h_2.$$

2. Лучи O_1h_1 и O_2h_2 лежат на параллельных прямых. Проведем плоскость через эти параллельные прямые. Через точки O_1 и O_2 про-

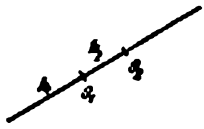


Рис. 5.2

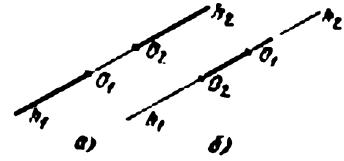


Рис. 5.3

ведем прямую, разбивающую плоскость на две полуплоскости с границей O_1O_2. Если оба луча лежат в одной из этих полуплоскостей (рис. 5.4), то такие лучи также называют *сонаправленными*. Если же лучи O_1h_1 и O_2h_2 лежат в различных полуплоскостях, то их на-

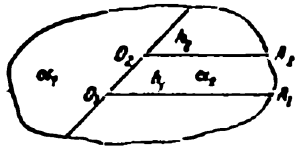


Рис. 5.4

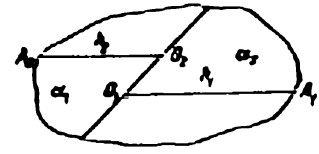


Рис. 5.5

зывают *противоположно направленными* (рис. 5.5). Для обозначения сонаправленных и противоположно направленных лучей в случае 2 используются те же обозначения, что и в случае 1. Для лучей, принадлежащих непараллельным прямым, понятие сонаправленности (или противоположной направленности) не вводится.

Множество всех лучей плоскости, каждый из которых сонаправлен с одним и тем же лучом, называется *направлением* на плоскости.

Множество всех лучей пространства, каждый из которых сонаправлен с одним и тем же лучом, называется *направлением* в пространстве.

*) Луч также обозначается одной малой латинской буквой (когда точка начала известна) или двумя заглавными буквами OA, где O — начало луча, а A — произвольная точка (открытого) луча.

1.2. Отрезок. Пусть A и B — две различные точки прямой a. Множество, состоящее из всех точек прямой a, лежащих между точками A и B, включая точки A, B, называется *отрезком*. Точки A и B называются *концами* отрезка, а все остальные точки — *внутренними точками* отрезка. Отрезок с концами A, B обозначается AB.

Длиной отрезка называют расстояние между его концами, измеренное каким-либо отрезком, принятым за единицу длины.

Длина отрезка AB часто обозначается $|AB|$. Отрезки считаются равными, если равны их длины.

§ 2. Углы на плоскости

2.1. Понятие угла. Пара различных лучей Oa и Ob, выходящих из одной точки O, называется *углом* и обозначается символом $\angle(a, b)$. Точка O называется *вершиной* угла, а лучи Oa и Ob — *сторонами* угла. Если A и B — две точки лучей Oa и Ob, то $\angle(a, b)$ обозначается также символом $\angle AOB$ (рис. 5.6).

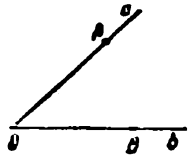


Рис. 5.6

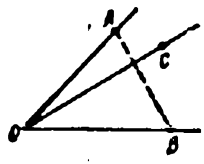


Рис. 5.7

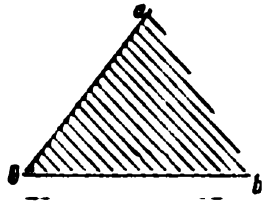


Рис. 5.8

Угол $\angle(a, b)$ называют *развернутым*, если лучи Oa и Ob, выходящие из одной точки, лежат на одной прямой и не совпадают (т. е. противоположно направлены).

Два угла считаются *равными*, если один угол можно наложить на другой так, чтобы стороны углов совпадали. *Биссектрисой* угла называется луч с началом в вершине угла, делящий угол на два равных угла.

Говорят, что луч, исходящий из вершины угла AOB, лежит между его сторонами, если он пересекает отрезок AB (рис. 5.7). Говорят, что точка C лежит между сторонами угла, если через эту точку можно провести луч с началом в вершине угла, лежащий между сторонами угла. Множество всех точек плоскости, лежащих между сторонами угла, образует *внутреннюю область* угла (рис. 5.8) *). Множество точек плоскости, не принадлежащих внутренней области и сторонам угла, образует *внешнюю область* угла.

Угол $\angle(a, b)$ считают больше угла $\angle(c, d)$, если угол $\angle(c, d)$ можно наложить на угол $\angle(a, b)$ так, что после совмещения одной пары сторон вторая сторона угла $\angle(c, d)$ будет лежать между сторонами угла $\angle(a, b)$. На рис. 5.9 $\angle AOB$ больше $\angle AOC$.

Пусть луч c лежит между сторонами угла $\angle(a, b)$ (рис. 5.10). Пары лучей a, c и c, b образуют два угла. Об угле $\angle(a, b)$ говорят, что он является *суммой* двух углов $\angle(a, c)$ и $\angle(c, b)$, и пишут

$$\angle(a, b) = \angle(a, c) + \angle(c, b).$$

*) Иногда углом также называют часть плоскости — объединение сторон угла и его внутренней области.

Обычно в геометрии имеют дело с углами, меньшими развернутого. Однако в результате сложения двух углов может получиться угол, больший развернутого. В этом случае часть плоскости, которая

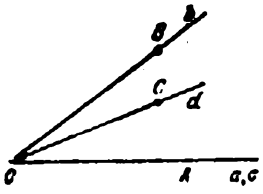


Рис. 5.9

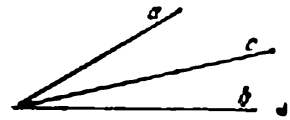


Рис. 5.10

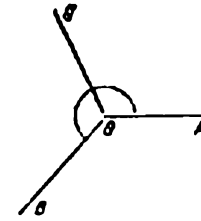


Рис. 5.11

считается внутренней областью угла, отмечают дугой. На рис. 5.11 внутренняя часть угла $\angle AOB$, полученного в результате сложения углов $\angle AOC$ и $\angle COB$ и большего развернутого, отмечена дугой.

2.2. Градусная мера измерения углов. При градусном измерении углов в качестве основной единицы измерения (эталонного угла, с которым сравниваются различные углы) берется угол в один градус (обозначается 1°). Угол в один *градус* — это угол, равный 1/180 части развернутого угла. Угол, равный 1/60 части угла в 1°,— это угол в одну *минуту* (обозначается 1′). Угол, равный 1/60 части угла в одну минуту,— это угол в одну *секунду* (обозначается 1″).

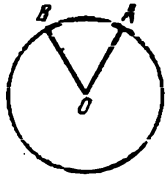


Рис. 5.12

2.3. Радианная мера измерения углов. Наряду с градусной мерой измерения углов в геометрии и тригонометрии употребляется и другая мера измерения углов, называемая *радианной*. Рассмотрим окружность радиуса R с центром O. Проведем два радиуса OA и OB так, чтобы длина дуги AB была равна радиусу окружности (рис. 5.12). Получившийся при этом центральный угол $\angle AOB$ будет углом в один *радиан*. Угол в 1 радиан принимается за единицу измерения радианной меры измерения углов. При радианном измерении углов развернутый угол равен π радиан.

Градусная и радианная единицы измерения углов связаны равенствами:

$$1 \text{ радиан} = \frac{180^\circ}{\pi} \approx 57^\circ 17' 45'';$$

$$1^\circ = \frac{\pi}{180} \text{ радиана} \approx 0{,}017453 \text{ радиана};$$

$$1' = \frac{\pi}{180 \cdot 60} \text{ радиана} \approx 0{,}000291 \text{ радиана};$$

$$1'' = \frac{\pi}{180 \cdot 60 \cdot 60} \text{ радиана} \approx 0{,}000005 \text{ радиана}.$$

Градусную (или радианную) меру угла называют также *величиной угла*. Величину угла $\angle AOB$ иногда обозначают $\widehat{AOB}$.

2.4. Классификация углов. Угол, равный 90°, или в радианной мере $\pi/2$, называется *прямым* углом; его часто обозначают буквой d. Угол, меньший 90°, называется *острым*; угол, больший 90°, но меньший 180°, называется *тупым*.

Два угла, имеющие одну общую сторону и в сумме составляющие 180°, называются *смежными* углами. Два угла, имеющие одну общую сторону и в сумме составляющие 90°, называются *дополнительными* углами.

2.5. Угол между направлениями. Если на плоскости выбраны два направления, то в каждой точке O плоскости начинается один луч OA первого направления и один луч OB второго направления. *Углом* между двумя направлениями на плоскости называется угол между лучами OA и OB. Угол между направлениями не зависит от выбора начальной точки лучей.

Углом между двумя направлениями в пространстве называют угол между двумя лучами этих направлений, имеющими общее начало.

§ 3. Параллельность и перпендикулярность на плоскости

3.1. Параллельность на плоскости. Две различные прямые a и b, лежащие в одной плоскости, называются *параллельными*, если они не имеют ни одной общей точки. Если прямые a и b параллельны, то пишут $a \parallel b$.

Множество всех прямых на плоскости, параллельных данной прямой, называют *пучком* параллельных прямых.

В результате пересечения двух прямых третьей образуются восемь углов (рис. 5.13). Пары углов $\angle 1$, $\angle 5$; $\angle 2$, $\angle 6$; $\angle 3$, $\angle 7$; $\angle 4$, $\angle 8$ называют *соответственными* углами; пары $\angle 3$, $\angle 5$; $\angle 4$, $\angle 6$ — *внутренними накрест лежащими* углами; пары $\angle 1$, $\angle 7$; $\angle 2$, $\angle 8$ — *внешними накрест лежащими* углами; пары $\angle 4$, $\angle 5$; $\angle 3$, $\angle 6$ — *прилежащими* углами.

Признаки параллельности прямых:

1) если при пересечении двух прямых a и b третьей прямой соответственные углы равны, то прямые a и b параллельны;

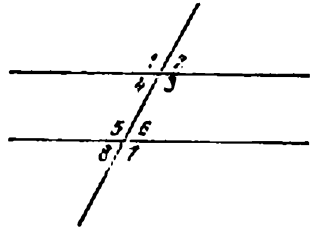


Рис. 5.13

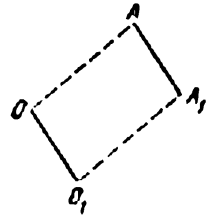


Рис. 5.14

2) если при пересечении двух прямых a и b третьей прямой внутренние (или внешние) накрест лежащие углы равны, то прямые a, b параллельны;

3) если при пересечении двух прямых a и b третьей прямой прилежащие углы в сумме составляют 180°, то прямые a и b параллельны.

Теоремы о равных отрезках.

1. Если от точек O и O_1 в одном и том же направлении отложены равные отрезки OA и O_1A_1, то отрезки OO_1 и AA_1 равны и параллельны *) (рис. 5.14).

*) *Параллельными отрезками* называют отрезки, принадлежащие параллельным прямым.

2. Теорема Фалеса. Если на одной прямой отложить несколько равных отрезков и через их концы провести параллельные прямые, пересекающие вторую прямую, то они отсекут на второй прямой равные отрезки (рис. 5.15).

Отрезки называют *пропорциональными*, если пропорциональны их длины.

Теоремы о пропорциональных отрезках.

1. Параллельные прямые, пересекающие стороны угла, отсекают на них пропорциональные отрезки.

На рис. 5.16 прямые AA_1 и BB_1 параллельны и

$$\frac{|OA_1|}{|OB_1|} = \frac{|OA|}{|OB|}.$$

2. Если отрезки OA_1 и OB_1 пропорциональны отрезкам OA и OB и лежат соответственно на лучах OA_1 и OA, то прямые AA_1 и BB_1 параллельны (рис. 5.16).

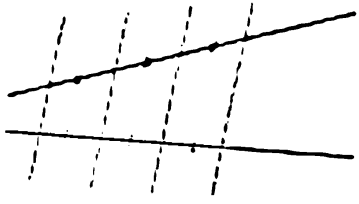


Рис. 5.15

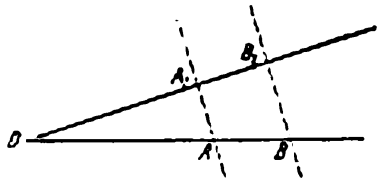


Рис. 5.16

3.2. Перпендикулярность на плоскости. Две прямые, при пересечении которых образуются прямые углы, называются *взаимно перпендикулярными*. Если прямые a и b взаимно перпендикулярны, то пишут

$$a \perp b.$$

Теоремы о перпендикулярных прямых.

1. Если на плоскости даны точка и прямая, то существует единственная прямая, проходящая через данную точку и перпендикулярная данной прямой.

2. Если прямая перпендикулярна одной из прямых пучка параллельных прямых, то она перпендикулярна любой другой прямой этого пучка.

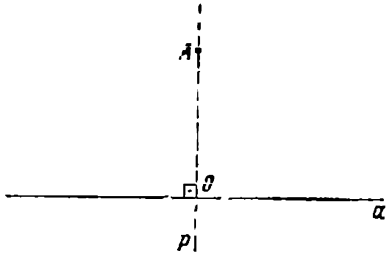


Рис. 5.17

3.3. Расстояние от точки до прямой. Пусть точка A лежит вне прямой a. Построим прямую p, перпендикулярную прямой a и проходящую через A (рис. 5.17). Обозначим через O точку пересечения прямых a и p. Точка O называется *основанием* перпендикуляра p.

Расстоянием от точки A до прямой a называется длина отрезка OA. Расстояние от точки A до любой точки прямой a, отличной от точки O, больше расстояния от точки A до прямой a.

§ 4. Параллельность и перпендикулярность в пространстве

4.1. Параллельность прямой и плоскости. Плоскость α и прямая a, не принадлежащая плоскости α, называются *параллельными*, если они не имеют ни одной общей точки. Любая прямая a, принадлежащая плоскости α, считается параллельной плоскости α. Если плоскость α и прямая a параллельны, то пишут $\alpha \parallel a$ или $a \parallel \alpha$.

Признак параллельности прямой и плоскости: если прямая параллельна какой-либо прямой, лежащей в плоскости, то данные прямая и плоскость параллельны.

Теоремы о плоскости и прямой, параллельной плоскости.

1. Если плоскость проведена через прямую, параллельную другой плоскости, и пересекает эту плоскость, то линия пересечения плоскостей параллельна данной прямой.

2. Если через каждую из двух параллельных прямых проведена произвольная плоскость и эти плоскости пересекаются, то линия их пересечения параллельна каждой из данных прямых.

4.2. Параллельность плоскостей. Две плоскости α и β называются *параллельными*, если они не имеют общей точки и не совпадают. Если плоскости α и β параллельны, то пишут $\alpha \parallel \beta$.

Признак параллельности двух плоскостей: Если две пересекающиеся прямые одной плоскости соответственно параллельны двум прямым другой плоскости, то эти плоскости параллельны.

Теоремы о параллельных плоскостях.

1. Если две параллельные плоскости α и β пересечены третьей плоскостью γ, то линии пересечения плоскостей α, γ и β, γ параллельны.

2. Через данную точку, не принадлежащую данной плоскости, можно провести одну и только одну плоскость, параллельную данной плоскости.

3. Если каждая из двух данных плоскостей параллельна третьей плоскости, то данные две плоскости параллельны между собой.

4.3. Перпендикулярность прямой и плоскости. Две прямые в пространстве называются *перпендикулярными*, если угол между направлениями, задаваемыми этими прямыми, равен 90°. Прямая и плоскость называются *взаимно перпендикулярными*, если прямая перпендикулярна каждой прямой, принадлежащей плоскости.

Если плоскость α перпендикулярна прямой a, то пишут $a \perp \alpha$ или $\alpha \perp a$.

Прямую, перпендикулярную плоскости, называют *перпендикуляром* к этой плоскости. Через каждую точку пространства можно провести одну и только одну прямую, перпендикулярную данной плоскости.

Признак перпендикулярности прямой и плоскости: если прямая a перпендикулярна каждой из двух пересекающихся прямых, лежащих в плоскости α, то прямая a и плоскость α взаимно перпендикулярны.

Теоремы о перпендикулярности прямой и плоскости.

1. Два различных перпендикуляра к плоскости параллельны.

2. Если одна из двух параллельных прямых перпендикулярна плоскости, то и другая перпендикулярна этой плоскости.

3. Прямая, перпендикулярная одной из двух параллельных плоскостей, перпендикулярна и другой плоскости.

4. Две плоскости, перпендикулярные одной и той же прямой, параллельны.

4.4. Расстояние от точки до плоскости. Пусть точка A_1 не принадлежит плоскости α. Через точку A_1 проведем перпендикуляр к плоскости. Точку пересечения перпендикуляра с плоскостью обозначим буквой A (рис. 5.18). Точка A есть *основание* перпендикуляра, а длина отрезка AA_1 — *расстояние от точки A_1 до плоскости α.*

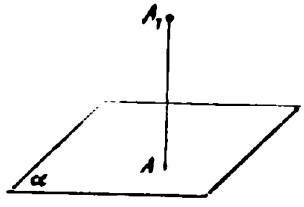


Рис 5.18

Расстояние от точки A_1 до плоскости α меньше расстояния от точки A_1 до любой точки плоскости α, отличной от точки A.

4.5. Перпендикулярность плоскостей. Две плоскости называются *взаимно перпендикулярными*, если угол между ними прямой *). Если плоскости α и β взаимно перпендикулярны, то пишут $\alpha \perp \beta$.

Признак перпендикулярности плоскостей: если плоскость содержит перпендикуляр к другой плоскости, то она перпендикулярна этой плоскости.

Теорема о взаимно перпендикулярных плоскостях.

Если две плоскости взаимно перпендикулярны, то прямая, принадлежащая одной плоскости и перпендикулярная линии пересечения плоскостей, перпендикулярна другой плоскости.

4.6. Наклонная. Прямую, пересекающую плоскость, но не перпендикулярную ей, называют *наклонной* к плоскости.

Теорема о трех перпендикулярах.

Для того чтобы прямая, лежащая в плоскости, была перпендикулярна наклонной, необходимо и достаточно, чтобы эта прямая была перпендикулярна проекции наклонной на эту плоскость.

4.7. Скрещивающиеся прямые. Две прямые в пространстве называются *скрещивающимися*, если они не пересекаются и не параллельны. Для обозначения пары скрещивающихся прямых a и b используется запись $a \div b$.

Признак скрещивающихся прямых: если одна из двух прямых лежит в некоторой плоскости, а другая прямая пересекает эту плоскость в точке, не принадлежащей первой прямой, то прямые скрещиваются.

Общим перпендикуляром двух скрещивающихся прямых a и b называется отрезок, удовлетворяющий условиям:

1) один конец отрезка принадлежит прямой a, а другой конец — прямой b;

2) прямая, содержащая отрезок, перпендикулярна как прямой a, так и прямой b.

Расстоянием между скрещивающимися прямыми называют длину общего перпендикуляра двух скрещивающихся прямых.

*) Об угле между плоскостями см. п. 6.3.

§ 5. Проектирование на плоскость

5.1. Параллельное проектирование. Пусть даны плоскость α и и прямая l, пересекающая плоскость α (рис. 5.19). Возьмем произвольную точку пространства A_1 и проведем через эту точку прямую l_1, параллельную l. Прямая l_1 пересечет плоскость α в некоторой точке A. Полученная таким образом точка A называется *проекцией* точки A_1 на плоскость α при проектировании параллельно прямой l. Обычно кратко говорят, что точка A есть *параллельная проекция* точки A_1.

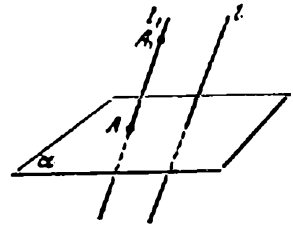


Рис. 5.19

Параллельной проекцией пространственной фигуры Φ_1 называется множество Φ параллельных проекций всех точек данной фигуры.

Свойства параллельного проектирования *):

1) проекция прямой есть прямая;

2) проекции параллельных прямых параллельны;

3) отношение проекций двух параллельных отрезков равно отношению проектируемых отрезков.

5.2. Ортогональное проектирование. Частным случаем параллельного проектирования является ортогональное проектирование.

Пусть даны плоскость α и прямая l, перпендикулярная α. Возьмем произвольную точку пространства A_1 и проведем через нее прямую l_1, параллельную l (и, следовательно, перпендикулярную плоскости α). Прямая l_1 пересечет плоскость α в некоторой точке A (рис. 5.20). Полученная точка A называется *ортогональной проекцией* точки A_1 на плоскость α.

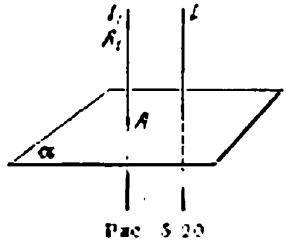


Рис. 5.20

Ортогональной проекцией фигуры Φ_1 на плоскость α называется множество Φ ортогональных проекций всех точек данной фигуры Φ_1. Как частный случай параллельного проектирования ортогональное проектирование обладает всеми свойствами параллельного проектирования.

Свойство ортогональной проекции плоского многоугольника.

Площадь s ортогональной проекции плоского многоугольника на плоскость α равна площади S проектируемого многоугольника, умноженной на косинус угла γ между плоскостью многоугольника и плоскостью α:

$$s = S \cdot \cos \gamma.$$

§ 6. Углы в пространстве

6.1. Угол между наклонной и плоскостью. *Углом между наклонной и плоскостью* называется угол между наклонной и ее ортогональной проекцией на плоскость. Угол между наклонной a и плоскостью α часто обозначается

$$\angle (a, \alpha).$$

*) Здесь предполагается, что проектирование производится параллельно прямой l, не параллельной проектируемым прямым или отрезкам.

6.2. Двугранный угол. Пара различных полуплоскостей α и β, имеющих общую границу — прямую a, называется *двугранным углом* (рис. 5.21). Прямая a называется *ребром* двугранного угла, а полуплоскости α и β — его *гранями*. Двугранный угол разбивает пространство на две части — *внутреннюю* и *внешнюю* области двугранного угла *). Грани двугранного угла считаются не принадлежащими ни той, ни другой области.

Двугранный угол обозначается символом $\angle$ и буквами, указывающими его грани и ребро. При этом буква, обозначающая ребро, ставится между буквами, обозначающими грани. Например, угол,

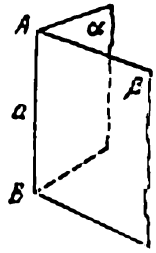


Рис. 5.21

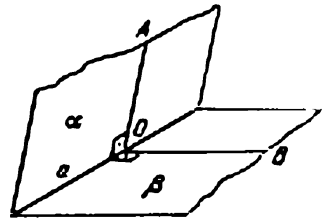


Рис. 5.22

изображенный на рис. 5.21, обозначается $\angle \alpha a \beta$. Применяется и краткое обозначение двугранного угла по его ребру, например: $\angle a$: $\angle AB$ (рис. 5.21).

При пересечении двугранного угла плоскостью, перпендикулярной его ребру, получается угол, называемый *линейным углом* двугранного угла. На рис. 5.22 $\angle AOB$ — линейный угол двугранного угла $\alpha a \beta$.

Двугранный угол измеряется своим линейным углом.

Двугранные углы равны, если равны их линейные углы.

6.3. Угол между двумя плоскостями. Две пересекающиеся плоскости определяют в пространстве четыре двугранных угла. Эти двугранные углы попарно равны, и сумма двух углов, имеющих общую грань, равна 180°. Меньший из двух углов называют *углом* между этими плоскостями. Если две плоскости параллельны, то угол между ними считается равным 0°.

Угол между плоскостями α, β обозначается $\angle (\alpha, \beta)$.

§ 7. Ломаная. Многоугольник

Ломаной линией (или просто *ломаной*) называется объединение отрезков, в котором конец каждого отрезка (кроме, быть может, последнего) является началом следующего отрезка, причем отрезки, имеющие общий конец, не лежат на одной прямой. Отрезки, составляющие ломаную, называют *звеньями* ломаной, а отрезки, имеющие общий конец,— *смежными звеньями* ломаной.

Ломаная называется *замкнутой*, если конец ее последнего звена совпадает с началом первого звена.

Ломаная называется *простой*, если каждое ее звено имеет только одну общую точку с другим звеном, и эта точка является концом

*) Иногда двугранным углом называется часть пространства — объединение внутренней области двугранного угла и его граней.

звена. На рис. 5.23 изображены простые ломаные. Ломаные, изображенные на рис. 5.24, не являются простыми.

Простая замкнутая ломаная разбивает плоскость на две области — внутреннюю и внешнюю относительно данной ломаной. Внешняя область обладает тем свойством, что в ней можно провести прямую, целиком принадлежащую области.

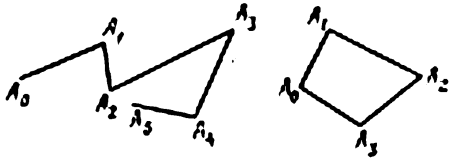


Рис. 5.23

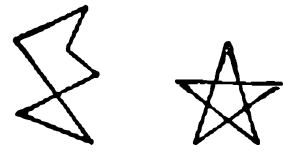


Рис. 5.24

Простая замкнутая ломаная вместе со своей внутренней областью называется *многоугольником* *). При этом сама ломаная называется *границей* многоугольника, а ее внутренняя область — *внутренней областью* многоугольника. Звенья границы многоугольника называются *сторонами* многоугольника, а точки пересечения звеньев — *вершинами* многоугольника. Число вершин многоугольника равно числу его сторон. Две стороны многоугольника, имеющие общую точку (вершину

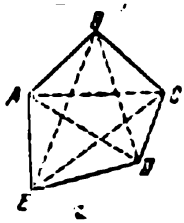


Рис. 5.25

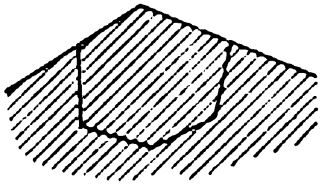


Рис. 5.26

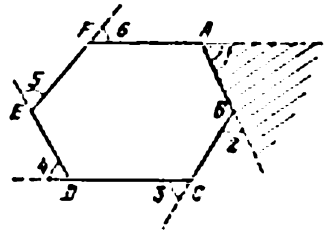


Рис. 5.27

многоугольника), называются *смежными сторонами*. Обычно многоугольник обозначается перечислением его вершин. На рис. 5.25 в многоугольнике $ABCDE$ отрезки AB, BC, CD, DE и AE — стороны многоугольника, точки A, B, C, D, E — вершины многоугольника.

Многоугольник называется *выпуклым*, если он целиком содержит отрезок, соединяющий две любые его точки, или, иначе, многоугольник называется *выпуклым*, если при продолжении любой из его сторон весь многоугольник лежит по одну сторону от этой прямой.

Проведем из вершины многоугольника два луча, содержащих две смежные стороны выпуклого многоугольника (рис. 5.26). Эти два луча разобьют плоскость на две области. Выпуклый многоугольник будет целиком находиться в одной из областей. Эту область называют *внутренним углом* (или просто *углом*) многоугольника. *Внешним углом* выпуклого многоугольника называется угол, смежный с его внутренним углом. На рис. 5.27 штриховкой отмечен один из внешних углов шестиугольника $ABCDEF$.

*) Иногда многоугольником также называется сама простая замкнутая ломаная.

Многоугольник именуется по числу своих углов. Так, многоугольник с тремя углами называется *треугольником*, многоугольник с четырьмя углами — *четырехугольником* и т. д.

Сумма внутренних углов выпуклого n-угольника равна $\pi(n-2)$.

Сумма внешних углов выпуклого многоугольника, взятых по одному при каждой вершине, равна 2π.

Сумма всех сторон многоугольника называется *периметром* многоугольника. Периметр обычно обозначается символом P_n, где индекс n указывает число сторон многоугольника.

Отрезок, соединяющий две не соседние вершины многоугольника, называется *диагональю* многоугольника. На рис. 5.25 AC, AD, BD, BE, CE — диагонали пятиугольника $ABCDE$. Число диагоналей выпуклого n-угольника равно $\frac{1}{2}n(n-3)$.

Многоугольник называется *правильным*, если все его стороны равны и все внутренние углы также равны между собой.

§ 8. Треугольники

8.1. Основные свойства. *Треугольником* называется многоугольник с тремя углами (и с тремя сторонами).

Стороны и углы треугольника считаются основными элементами треугольника.

Говорят, что в треугольнике ABC (рис. 5.28) сторона a лежит против угла α и, наоборот, против стороны a лежит угол α. Аналогично b лежит против β, c — против γ.

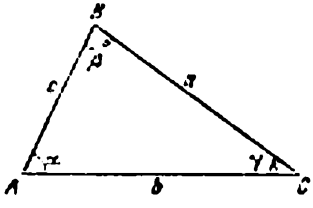


Рис. 5.28

Треугольник полностью определяется *) любой из следующих троек своих основных элементов: либо тремя сторонами, либо одной стороной и двумя одинаково расположенными углами, либо двумя сторонами и углом между ними.

Для существования треугольника, задаваемого тремя сторонами a, b, c, необходимо и достаточно выполнение неравенств, называемых *неравенствами треугольника*:

$$a+b>c, \quad a+c>b, \quad b+c>a.$$

Для существования треугольника, задаваемого стороной a и углами α, β, необходимо и достаточно выполнение неравенства

$$\alpha+\beta<180^\circ,$$

а для существования треугольника, задаваемого сторонами b, c и углом γ между ними, необходимо и достаточно выполнение неравенства

$$\gamma<180^\circ.$$

В качестве тройки элементов, однозначно определяющих треугольник, можно выбирать и другие наборы элементов. Некоторые из таких наборов приведены в § 13.

*) Точнее говоря, задание перечисленных ниже троек элементов определяет множество равных треугольников. Равенство этих троек элементов для двух треугольников также называют *признаками равенства* треугольников.

Не всякая тройка основных элементов треугольника однозначно задает треугольник. Так, задавая три угла треугольника α, β, γ (они не являются независимыми и связаны между собой равенством $\alpha+\beta+\gamma=180°$), можно построить сколь угодно много не равных треугольников с углами α, β, γ (эти треугольники подобны).

Соотношения между сторонами и углами треугольника:

1) против большей стороны лежит больший угол;

2) против большего угла лежит большая сторона;

3) против равных сторон лежат равные углы, и, обратно, против равных углов лежат равные стороны.

Соотношение между внутренними и внешними углами треугольника: сумма двух любых внутренних углов треугольника равна внешнему углу треугольника, смежного с третьим углом.

Стороны и углы треугольника связаны между собой также соотношениями, называемыми теоремой синусов и теоремой косинусов (см. гл. 6).

Треугольник называется *тупоугольным*, *прямоугольным* или *остроугольным*, если его наибольший внутренний угол соответственно больше, равен или меньше 90°.

Площадь треугольника S может быть вычислена по формулам:

$$S=\frac{1}{2}ah_a=\frac{1}{2}bh_b=\frac{1}{2}ch_c,$$

$$S=\sqrt{p(p-a)(p-b)(p-c)}\quad (\textit{формула Герона}),$$

$$S=\frac{1}{2}ab\sin\gamma,\quad S=\frac{abc}{4R},\quad S=pr,$$

Здесь a, b, c — стороны треугольника; h_a, h_b, h_c — высоты, опущенные на стороны a, b и c соответственно; $p=\frac{1}{2}(a+b+c)$ — полупериметр, R — радиус окружности, описанной около треугольника, r — радиус окружности, вписанной в треугольник.

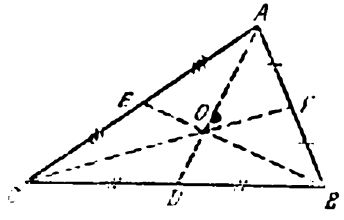


Рис. 5.29

Два треугольника называются *равновеликими*, если их площади равны.

8.2. Медианы треугольника. *Медианой* треугольника называется отрезок, соединяющий вершину треугольника с серединой противоположной стороны (на рис. 5.29 отрезки AD, BE, CF — медианы треугольника ABC). Медианы пересекаются в одной точке, лежащей внутри треугольника.

Основные свойства медиан треугольника:

1) медианы треугольника точкой их пересечения делятся в отношении 2:1 (считая от вершин треугольника);

2) медиана делит треугольник на два равновеликих треугольника;

3) три медианы треугольника делят треугольник на шесть равновеликих треугольников (на рис. 5,29 треугольники AOE, EOC, COD, DOB, BOF, FOA равновелики).

Медиана треугольника m_a, проведенная к стороне a, выражается через стороны треугольника по формуле

$$m_a = \frac{1}{2}\sqrt{2b^2 + 2c^2 - a^2}.$$

8.3. Высоты треугольника. Пусть ABC — произвольный треугольник. Проведем через вершину A перпендикуляр к прямой a, содержащей сторону BC (рис. 5.30). Обозначим основание перпендикуляра буквой D. Отрезок перпендикуляра AD называют *высотой* треугольника ABC, опущенной из вершины A на сторону BC. Сторону BC при этом называют *основанием* треугольника ABC.

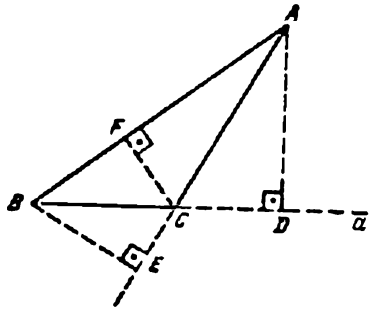


Рис. 5.30

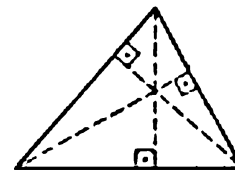


Рис. 5.31

В остроугольном треугольнике (рис. 5.31) все три высоты лежат внутри треугольника. В тупоугольном треугольнике ABC (см. рис. 5.30) две высоты (AD и BE) пересекают продолжение сторон и лежат вне треугольника; третья высота (CF) пересекает сторону треугольника. В прямоугольном треугольнике катеты являются также и высотами. Три прямые, содержащие разные высоты треугольника, всегда пересекаются в одной точке, называемой *ортоцентром* треугольника. В остроугольном треугольнике ортоцентр лежит внутри треугольника; в тупоугольном — вне треугольника; в прямоугольном треугольнике ортоцентр совпадает с вершиной прямого угла. Высоты треугольника, опущенные на стороны треугольника a, b и c, обозначаются h_a, h_b и h_c соответственно.

Высота треугольника h_a выражается через стороны по формуле

$$h_a = \frac{2\sqrt{p(p-a)(p-b)(p-c)}}{a},$$

где $p = \frac{1}{2}(a+b+c)$.

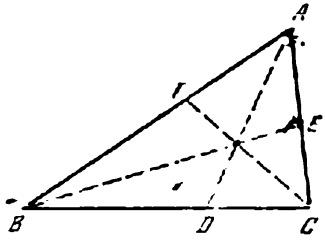


Рис. 5.32

8.4. Биссектрисы треугольника. Пусть ABC — произвольный треугольник. Проведем биссектрису угла BAC; она пересекает сторону BC в точке D (рис. 5.32). Отрезок биссектрисы внутреннего угла треугольника называют *биссектрисой* треугольника. Три биссектрисы треугольника (AD, BE, CF на рис. 5.32) пересекаются в одной точке, лежащей внутри треугольника и являющейся центром вписанной в треугольник окружности.

Свойства биссектрисы угла треугольника:

1) биссектриса делит противоположную сторону на части, пропорциональные прилежащим сторонам. Так, для треугольника ABC,

изображенного на рис. 5.32,

$$\frac{|AE|}{|EC|}=\frac{|AB|}{|BC|};$$

2) биссектриса треугольника делит площадь треугольника в отношении, пропорциональном прилежащим сторонам:

$$\frac{S_{\triangle ABE}}{S_{\triangle BCE}}=\frac{|AB|}{|BC|}.$$

Биссектриса треугольника l_a, делящая пополам угол α, противолежащий стороне a, вычисляется через стороны и углы треугольника по формулам

$$l_a=\frac{2ab\cos\frac{\gamma}{2}}{a+b}=\frac{2ac\cos\frac{\beta}{2}}{a+c},$$

где b и c — стороны треугольника, лежащие против углов β и γ соответственно.

8.5. Средняя линия треугольника. *Средней линией* треугольника называется отрезок, соединяющий середины двух сторон треугольника (на рис. 5.33 DE, DF, FE — средние линии).

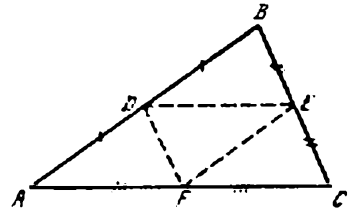


Рис. 5.33

Свойства средней линии треугольника:

1) прямая, содержащая среднюю линию треугольника, параллельна прямой, содержащей третью сторону треугольника;

2) средняя линия треугольника равна половине его третьей стороны;

3) средняя линия треугольника отсекает от треугольника подобный треугольник. Площадь отсекаемого треугольника относится к площади основного треугольника как 1:4.

8.6. Равнобедренный треугольник. Треугольник, у которого две стороны равны, называется *равнобедренным*.

В равнобедренном треугольнике обычно за *основание* принимают сторону, не равную никакой из остальных двух сторон.

Вершину угла, лежащего против основания равнобедренного треугольника, обычно называют *вершиной* равнобедренного треугольника.

Свойства равнобедренного треугольника:

1) в равнобедренном треугольнике углы при основании треугольника равны;

2) высота, проведенная из вершины, является также биссектрисой и медианой.

8.7. Равносторонний треугольник. Треугольник, все стороны которого равны, называется *равносторонним* (или *правильным*) треугольником.

Свойства равностороннего треугольника:

1) все углы равны (каждый угол равен 60°);

2) каждая из трех высот является также биссектрисой и медианой;

3) центр окружности, описанной около треугольника, совпадает с центром окружности, вписанной в него.

8.8. Прямоугольный треугольник. Треугольник, один из углов которого прямой, называется *прямоугольным* треугольником. Сторона, лежащая против прямого угла, называется *гипотенузой*, а две остальные стороны — *катетами*. На рис. 5.34 изображен прямоугольный треугольник: $\widehat{BAC}=90°$, BC — гипотенуза, AB и AC — катеты.

Прямоугольный треугольник, имеющий равные катеты, называется *равнобедренным* прямоугольным треугольником. В равнобедренном

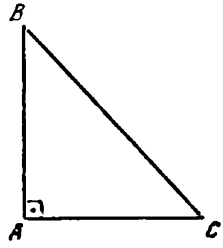


Рис. 5.34

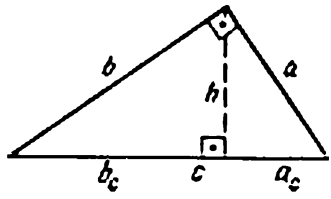


Рис. 5.35

прямоугольном треугольнике острые углы равны (каждый из них равен 45°).

Стороны всякого прямоугольного треугольника a, b и c (c — гипотенуза) связаны между собой соотношением, называемым теоремой Пифагора:

$$c^2=a^2+b^2,$$

которая читается так: квадрат гипотенузы равен сумме квадратов катетов.

Свойства прямоугольного треугольника:

1) катет есть среднее пропорциональное между гипотенузой и проекцией этого катета на гипотенузу (рис. 5.35):

$$b_c:b=b:c, \quad a_c:a=a:c,$$

или

$$b^2=b_cc, \qquad a^2=a_cc;$$

здесь через a_c и b_c обозначены проекции катетов a и b на гипотенузу c;

2) высота, проведенная из вершины прямого угла, есть среднее пропорциональное между проекциями катетов на гипотенузу (рис. 5.35):

$$b_c:h=h:a_c \quad \text{или} \quad h^2=a_cb_c;$$

3) центр окружности, описанной около прямоугольного треугольника, лежит на середине гипотенузы.

Площадь прямоугольного треугольника равна половине произведения его катетов:

$$S=\frac{1}{2}ab.$$

§ 9. Четырехугольники

9.1. Параллелограмм. Четырехугольник, противоположные стороны которого попарно параллельны, называется *параллелограммом* (рис. 5.36).

Свойства параллелограмма:

1) середина диагонали параллелограмма является его центром симметрии;

2) противоположные стороны параллелограмма равны;

3) противоположные углы параллелограмма равны;

4) каждая диагональ параллелограмма делит его на два равных треугольника;

5) диагонали параллелограмма делятся точкой пересечения пополам;

6) сумма квадратов диагоналей параллелограмма (d_1 и d_2) равна сумме квадратов его сторон:

$$d_1^2 + d_2^2 = 2(a^2 + b^2).$$

Признаки параллелограмма:

1) если в четырехугольнике противоположные стороны попарно равны, то этот четырехугольник — параллелограмм;

2) если в четырехугольнике две противоположные стороны равны и параллельны, то этот четырехугольник — параллелограмм.

Каждый из признаков параллелограмма может быть взят в качестве определения параллелограмма. Так, из первого признака получаем следующее определение параллелограмма.

Четырехугольник, у которого противоположные стороны попарно равны, называется *параллелограммом*.

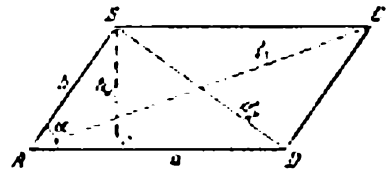


Рис. 5.36

Сумма внутренних углов параллелограмма равна 2π (360°).

Отрезок перпендикуляра к сторонам параллелограмма, заключенный между ними, называется *высотой* параллелограмма.

Площадь параллелограмма равна:

1) произведению его основания на высоту (рис. 5.36):

$$S = ah_a;$$

2) произведению смежных сторон параллелограмма на синус угла между ними (рис. 5.36):

$$S = ab \sin \alpha.$$

9.2. Ромб. Параллелограмм, все стороны которого равны, называется *ромбом* (рис. 5.37). Помимо всех свойств параллелограмма, ромб обладает следующими специальными свойствами:

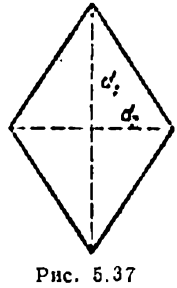


Рис. 5.37

1) прямая, содержащая диагональ ромба, является его осью симметрии;

2) диагонали ромба взаимно перпендикулярны;

3) диагонали ромба являются биссектрисами его внутренних углов.

Площадь ромба может быть вычислена по формуле

$$S = \frac{1}{2} d_1 d_2,$$

где d_1 и d_2 — диагонали ромба.

9.3. Прямоугольник. Параллелограмм, у которого все углы прямые, называется *прямоугольником* (рис. 5.38).

Помимо всех свойств параллелограмма, прямоугольник обладает следующими специальными свойствами:

1) перпендикуляр, проходящий через середины противоположных сторон прямоугольника, является его осью симметрии;

2) прямоугольник имеет две оси симметрии;

3) диагонали прямоугольника равны.

Площадь прямоугольника вычисляется по формуле

$$S = ab,$$

где a и b — смежные стороны прямоугольника.

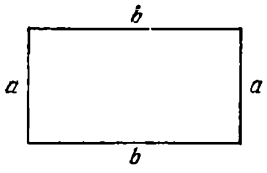


Рис. 5.38

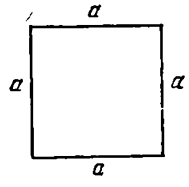


Рис. 5.39

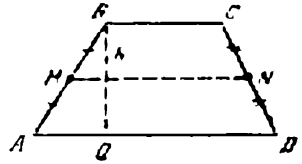


Рис. 5.40

9.4. Квадрат. Прямоугольник, у которого все стороны равны, называется *квадратом*.

Из определений квадрата и ромба следует, что квадрат (рис. 5.39) является ромбом, у которого все углы прямые. Так как квадрат является и параллелограммом, и прямоугольником, и ромбом, то все свойства этих фигур присущи и квадрату.

Площадь квадрата вычисляется по формуле

$$S = a^2,$$

где a — сторона квадрата.

9.5. Трапеция. Четырехугольник, две стороны которого параллельны, а две другие — непараллельны, называется *трапецией*.

Параллельные стороны трапеции называются ее *основаниями*, а непараллельные — *боковыми сторонами* (на рис. 5.40 BC и AD — основания, AB и CD — боковые стороны).

Отрезок перпендикуляра к основаниям трапеции, заключенный между основаниями, называется *высотой* трапеции (на рис. 5.40 отрезок BQ — высота). Трапеция, боковые стороны которой равны ($|AB| = |CD|$), называется *равнобочной* (или *равнобедренной*). В равнобочной трапеции углы при основании попарно равны:

$$\angle BAD = \angle ADC, \quad \angle ABC = \angle BCD.$$

Отрезок, соединяющий середины боковых сторон трапеции, называется *средней линией* трапеции. Средняя линия параллельна основаниям, равна их полусумме и делит высоту трапеции пополам.

Площадь трапеции вычисляется по формуле

$$S=\frac{a+b}{2}h,$$

где a и b — основания, h — высота трапеции.

§ 10. Подобные многоугольники

10.1. Признак подобия многоугольников. Если стороны одного многоугольника пропорциональны сторонам другого многоугольника и соответственные углы (т. е. углы, лежащие между пропорциональными сторонами) равны, то многоугольники *подобны*.

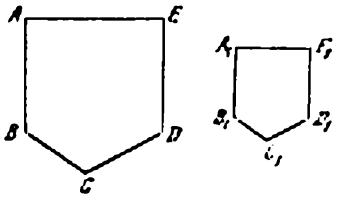


Рис. 5.41

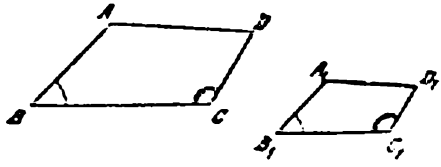


Рис. 5.42

На рис. 5.41 изображен пятиугольник $ABCDE$, подобный пятиугольнику $A_1B_1C_1D_1E_1$ с *коэффициентом подобия* $k=2$:

$$\frac{|AB|}{|A_1B_1|}=\frac{|BC|}{|B_1C_1|}=\frac{|CD|}{|C_1D_1|}=\frac{|DE|}{|D_1E_1|}=\frac{|AE|}{|A_1E_1|}=k=2.$$

Два четырехугольника *подобны*, если у них три пары соответственных сторон пропорциональны и пары соответственных углов, заключенных между соответственными сторонами, равны (рис. 5.42):

$$\frac{|AB|}{|A_1B_1|}=\frac{|BC}{|B_1C_1|}=\frac{|CD|}{|C_1D_1|},\quad \angle B=\angle B_1,\quad \angle C=\angle C_1.$$

Свойства подобных многоугольников:

1) отношение периметров подобных многоугольников равно отношению их соответственных сторон (коэффициенту подобия);

2) отношение площадей подобных многоугольников равно квадрату коэффициента подобия.

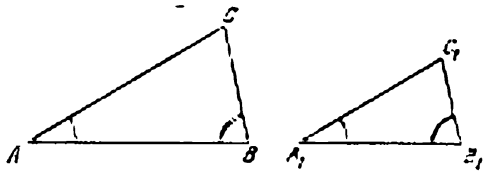


Рис. 5.43

10.2. Признаки подобия треугольников.

1. Если три стороны одного треугольника пропорциональны трем сторонам другого треугольника, то треугольники подобны (рис. 5.43):

$$\frac{|AB|}{|A_1B_1|}=\frac{|BC|}{|B_1C_1|}=\frac{|AC|}{|A_1C_1|}.$$

2. Если два угла одного треугольника равны двум углам другого треугольника, то треугольники подобны (рис. 5.43):

$$\angle A = \angle A_1, \quad \angle B = \angle B_1.$$

3. Если две стороны одного треугольника пропорциональны двум сторонам другого треугольника и углы между этими сторонами равны, то треугольники подобны (рис. 5.43):

$$\frac{|AB|}{|A_1B_1|} = \frac{|BC|}{|B_1C_1|}, \quad \angle B = \angle B_1.$$

Для подобия треугольников специального вида (прямоугольных, равнобедренных) необходимо выполнение меньшего числа условий, нежели для произвольных треугольников. Так, прямоугольные треугольники подобны, если либо гипотенуза и катет одного треугольника пропорциональны гипотенузе и катету другого треугольника, либо острый угол одного треугольника равен острому углу другого.

Свойства подобных треугольников:

1) отношение периметров подобных треугольников равно отношению их соответственных сторон (коэффициенту подобия);

2) отношение высот подобных треугольников, опущенных на соответственные стороны, равно отношению соответственных сторон (коэффициенту подобия);

3) отношение площадей подобных треугольников равно квадрату коэффициента подобия.

§ 11. Окружность и круг

11.1. Окружность и круг. *Окружностью* называется множество всех точек плоскости, находящихся на заданном отличном от нуля расстоянии от некоторой данной точки плоскости, называемой *центром* окружности.

Радиусом окружности называется отрезок, соединяющий центр окружности с любой точкой окружности (на рис. 5.44 отрезок OA — радиус). Радиус окружности обычно обозначается буквами r, R.

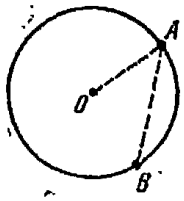


Рис. 5.44

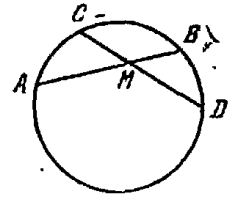


Рис. 5.45

Отрезок, соединяющий две точки окружности, называется *хордой* (на рис. 5.44 AB — хорда).

Хорда, проходящая через центр окружности, называется *диаметром*. Диаметр равен удвоенному радиусу окружности. Диаметр обычно обозначается буквами d, D.

Свойства хорд окружности:

1) диаметр, делящий хорду пополам, перпендикулярен этой хорде;

2) равные хорды равноудалены от центра окружности, и, наоборот, равноудаленные от центра окружности хорды равны;

3) из двух не равных хорд окружности бо́льшая хорда расположена ближе к центру окружности, и, наоборот, из двух не равных хорд большей будет та, которая ближе к центру;

4) отрезки пересекающихся хорд AM, MB, CM и MD (рис. 5.45) связаны равенством

$$|AM|\cdot|MB|=|CM|\cdot|MD|.$$

Кругом называется множество всех точек плоскости, расстояние которых от некоторой данной точки плоскости (*центра* круга) не больше данного. Граница круга — окружность.

Радиус, хорда, диаметр и центр окружности являются радиусом, хордой, диаметром и центром круга, ограниченного этой окружностью.

11.2. Касательная и секущая. Прямая, имеющая с окружностью только одну общую точку, называется *касательной* к этой окружности.

Для того чтобы прямая была касательной к окружности, необходимо и достаточно, чтобы эта прямая была перпендикулярна диаметру окружности и проходила через его конец.

Через любую точку, лежащую вне окружности, можно провести две различные касательные (рис. 5.46).

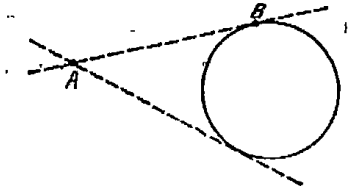


Рис. 5.46

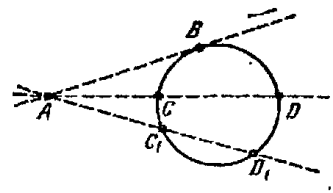


Рис. 5.47

Прямая, имеющая с окружностью две общие точки, называется *секущей*. На рис. 5.47 AD и AD_1 — секущие.

Если через точку A, лежащую вне круга, провести касательную и секущую (см. рис. 5.47), то отрезки касательной и секущей связаны равенством

$$|AB|^2=|AD|\cdot|AC|=|AD_1|\cdot|AC_1|.$$

11.3. Взаимное расположение двух окружностей. Пусть на плоскости даны две несовпадающие точки O_1 и O_2, расстояние между которыми $h=|O_1O_2|$, и две окружности с радиусами R_1 и R_2 и центрами в точках O_1 и O_2 соответственно. Для определенности положим $R_1\geqslant R_2$. При заданном расстоянии $|O_1O_2|=h$ и заданных радиусах R_1 и R_2 могут быть следующие случаи взаимного расположения окружностей:

1) если $h<R_1-R_2$, то окружности не пересекаются и круг радиуса R_2 целиком принадлежит кругу радиуса R_1 (рис. 5.48);

2) если $h=R_1-R_2$, то круг радиуса R_2 целиком принадлежит кругу радиуса R_1, а окружности имеют одну общую точку M (рис. 5.49). О таком расположении окружностей говорят, что они касаются друг друга изнутри в точке M;

3) если $h>R_1+R_2$, то окружности не пересекаются, а круги не имеют ни одной общей точки (рис. 5.50);

4) если $h=R_1+R_2$, то окружности (и круги) имеют одну общую точку M (рис. 5.51). О таком расположении окружностей говорят, что они касаются друг друга снаружи в точке M;

5) если $R_1 - R_2 < h < R_1 + R_2$, то окружности имеют две общие точки. В этом случае говорят, что окружности пересекаются в точках M_1 и M_2 (рис. 5.52).

Отрезок O_1O_2 называется *линией центров* окружностей. Прямая, касательная к обеим окружностям и не пересекающая линии центров,

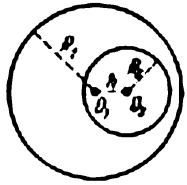


Рис. 5.48

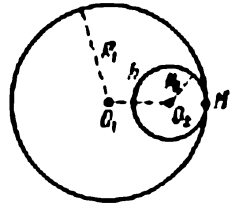


Рис. 5.49

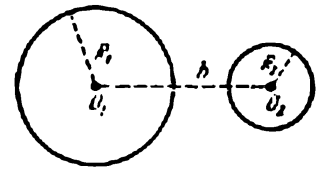


Рис. 5.50

называется *внешней касательной* двух данных окружностей. В перечисленных выше пяти случаях внешняя касательная к двум окружностям может быть построена во всех случаях, кроме 1).

Прямая, касающаяся обеих окружностей и пересекающая их линию центров, называется *внутренней касательной* двух данных

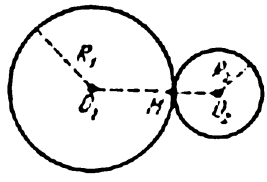


Рис. 5.51

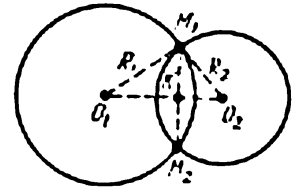


Рис. 5.52

окружностей. Внутренняя касательная может быть построена в случаях 3) и 4).

Если две окружности пересекаются (см. случай 5), рис. 5.52), то отрезок M_1M_2 называется *общей хордой* двух пересекающихся окружностей. Общая хорда двух пересекающихся окружностей взаимно перпендикулярна с линией центров и точкой пересечения делится на два равных отрезка:

$$|M_1K| = |KM_2|.$$

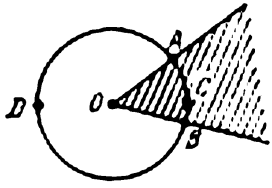


Рис. 5.53

11.4. Центральные углы и дуги окружности. Угол с вершиной в центре окружности называется *центральным* углом окружности.

Два различных луча, выходящих из центра окружности, определяют два центральных угла. Точки пересечения лучей с окружностью (на рис. 5.53 точки A и B) делят окружность на две части. Эти части окружности называются *дугами* окружности. Чтобы выделить одну из указанных дуг, на дугах выбирают произвольные точки (на рис. 5.53 точки C и D) и говорят о дугах ACB и ADB. Для обозначения дуг используется символ $\smile$. Так, например, дуги ACB

и ADB обозначаются

$$\smile ACB \quad \text{и} \quad \smile ADB.$$

Центрами дуг называют центр окружности.

Дуги окружности измеряют в градусах, минутах и секундах. Сколько градусов, минут и секунд имеет данный центральный угол, столько же градусов, минут и секунд имеет и соответствующая дуга. Угловая величина дуги ACB ($\smile ACB$) иногда обозначается символом $\overset{\frown}{ACB}$.

Дуга, соответствующая центральному углу в 180°, называется *полуокружностью*.

Между дугами и соответствующими им центральными углами окружности имеют место следующие отношения:

1) две дуги, принадлежащие окружностям одного и того же радиуса, равны в том и только в том случае, когда их угловые величины равны *);

2) в одной и той же окружности большему центральному углу соответствует бо́льшая дуга.

11.5. Дуги и хорды окружности. Пусть в окружности проведена хорда AB, не проходящая через центр окружности (рис. 5.54). О хорде AB говорят, что она стягивает дугу AB (при этом предполагается, что из двух дуг, на которые точки A и B разбивают окружность, выбирается наименьшая дуга, т. е. угловая величина дуги AB заключена в промежутке (0°, 180°)). Между хордами окружности и стягиваемыми ими дугами имеют место следующие отношения:

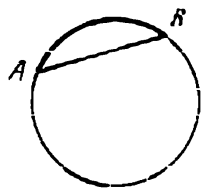


Рис. 5 54

1) равные дуги стягиваются равными хордами;

2) равные хорды стягивают равные дуги;

3) диаметр, перпендикулярный хорде, делит стягиваемую хордой дугу пополам.

11.6. Углы в окружности. Угол, вершина которого принадлежит окружности, а стороны пересекают окружность, называют *углом, вписанным* в эту окружность (рис. 5.55). Говорят также, что угол

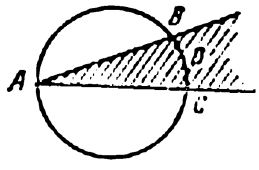


Рис. 5.55

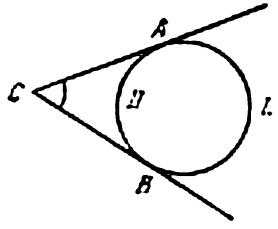


Рис. 5.56

BAC опирается на дугу BDC. Величина вписанного угла равна половине угловой величины дуги, на которую он опирается (или половине величины центрального угла, соответствующего данной дуге):

$$\overset{\frown}{BAC} = \frac{1}{2}\overset{\frown}{BDC}.$$

*) Угловые величины дуг окружностей разных радиусов могут быть равны, но дуги тем не менее не будут равны.

Угол, образованный двумя касательными к окружности, проходящими через одну точку, называется *описанным углом* (на рис. 5.56 CA и CB — касательные). Величина описанного угла равна полуразности угловых величин дуг, заключенных между его сторонами:

$$\widehat{ACB} = \frac{1}{2}(\overset{\smile}{ALB} - \overset{\smile}{ADB}).$$

Величина угла, образованного двумя секущими, имеющими общую точку, лежащую вне окружности, равна полуразности угловых величин дуг, заключенных между его сторонами (рис. 5.57):

$$\widehat{ACA_1} = \frac{1}{2}(\overset{\smile}{ADA_1} - \overset{\smile}{BLB_1}).$$

Величина угла, образованного двумя секущими, имеющими общую точку, лежащую внутри окружности, равна полусумме угловых

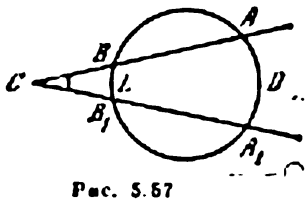


Рис. 5.57

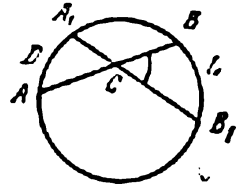


Рис. 5.58

величин дуг, заключенных между его сторонами (рис. 5.58):

$$\widehat{BCB_1} = \frac{1}{2}(\overset{\smile}{BLB_1} + \overset{\smile}{ADA_1}).$$

11.7. Длины и площади в окружности и круге. *Длиной окружности* называется предел последовательности периметров правильных многоугольников, вписанных в данную окружность, при неограниченном увеличении числа сторон *).

Длина окружности L вычисляется по формуле

$$L = \pi d,$$

где d — диаметр окружности, или по формуле

$$L = 2\pi r,$$

где r — радиус окружности.

Длина дуги окружности с угловой величиной α° вычисляется по формуле

$$l = \frac{\pi r \alpha}{180}.$$

Площадью круга называется предел последовательности площадей правильных многоугольников, вписанных в данную окружность, при неограниченном увеличении числа сторон.

Площадь круга радиуса r вычисляется по формуле

$$S = \pi r^2.$$

*) Вывод формул длины окружности и площади круга см. в п. 6.1 гл. 8.

Сектором называется часть круга, ограниченная двумя его радиусами (рис. 5.59).

Площадь сектора с угловой величиной дуги α° вычисляется по формуле

$$S_{\text{сект}} = \frac{\pi r^2 \alpha}{360}.$$

Часть круга, ограниченная хордой и стягиваемой ею дугой, называется *сегментом* (рис. 5.60).

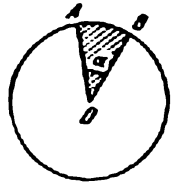


Рис. 5.59

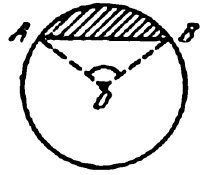


Рис. 5.60

Площадь сегмента вычисляется как разность площади сектора, ограниченного радиусами OA и OB, и площади треугольника AOB (см. рис. 5.60).

§ 12. Многоугольники и окружности

§ 12.1. Вписанные и описанные многоугольники. Многоугольник, все вершины которого принадлежат окружности, называется *вписанным* в эту окружность, а окружность — *описанной* около многоугольника.

Многоугольник, все стороны которого касаются окружности, называется *описанным* около этой окружности, а окружность — *вписанной* в этот многоугольник.

Около всякого правильного многоугольника можно описать окружность, и во всякий правильный многоугольник можно вписать окружность.

Центр вписанной в правильный многоугольник окружности совпадает с центром описанной около правильного многоугольника окружности; эта точка называется *центром* правильного многоугольника.

Отрезок перпендикуляра, проведенного из центра правильного многоугольника к его стороне, называется *апофемой* правильного многоугольника.

Сторона правильного n-угольника a_n выражается через радиус R описанной около него окружности и число сторон n по формуле

$$a_n = 2R \sin \frac{180^\circ}{n}.$$

Площадь правильного n-угольника равна половине произведения его периметра на радиус вписанной окружности:

$$S_n = \frac{1}{2} P_n r,$$

Площадь правильного n-угольника выражается через радиус описанной окружности R и число сторон n по формуле

$$S_n = \frac{1}{2} R^2 n \sin \frac{360^\circ}{n}.$$

12.2. Вписанные треугольники. Треугольник, все вершины которого принадлежат окружности, называется *вписанным* в эту окружность, а окружность — *описанной* около этого треугольника. Около всякого треугольника можно описать окружность, и притом только одну. Центр описанной около треугольника окружности — точка пересечения перпендикуляров, восставленных к сторонам этого треугольника и проходящих через их середины; он лежит внутри треугольника, если треугольник остроугольный; вне треугольника, если треугольник тупоугольный; на середине гипотенузы, если треугольник прямоугольный.

Радиус окружности, описанной около произвольного треугольника, вычисляется по формулам

$$R = \frac{1}{2} \cdot \frac{a}{\sin \alpha} = \frac{1}{2} \cdot \frac{b}{\sin \beta} = \frac{1}{2} \cdot \frac{c}{\sin \gamma},$$

$$R = \frac{abc}{4S_\triangle} = \frac{abc}{4\sqrt{p(p-a)(p-b)(p-c)}},$$

где a, b, c — стороны треугольника, $p = \frac{1}{2}(a+b+c)$ — полупериметр, $S_\triangle$ — площадь треугольника, α, β, γ — углы треугольника, лежащие против сторон a, b, c соответственно.

12.3. Описанные треугольники. Треугольник, все стороны которого касаются окружности, называется *описанным* около этой окружности, а окружность — *вписанной* в этот треугольник.

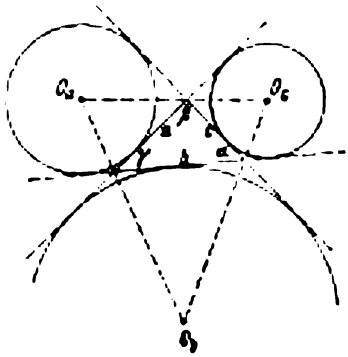


Рис. 5.61

Во всякий треугольник можно вписать окружность, и притом только одну. Центр O вписанной в треугольник окружности является точкой пересечения биссектрис внутренних углов треугольника.

Радиус окружности, вписанной в произвольный треугольник, вычисляется по формуле

$$r = \frac{S_\triangle}{p} = \sqrt{\frac{(p-a)(p-b)(p-c)}{p}}.$$

12.4. Вневписанная окружность. *Вневписанной* окружностью называется окружность, касающаяся одной стороны треугольника и продолжения двух других сторон.

Биссектрисы пары внешних углов треугольника, смежных с углами β и γ (рис. 5.61), пересекаются в точке O_a. Через эту же точку проходит биссектриса внутреннего угла α. Точка O_a — центр вневписанной окружности, касающейся стороны a и продолжения сторон b и c. Аналогично находятся точки O_b и O_c — центры вневписанных окружностей, касающихся сторон b и c соответственно (см. рис. 5,61).

Радиусы вневписанных окружностей r_a, r_b, r_c, касающихся сторон a, b и c соответственно, вычисляются по формулам

$$r_a = \frac{S_\triangle}{p-a} = \frac{2S_\triangle}{b+c-a},$$

$$r_b = \frac{S_\triangle}{p-b} = \frac{2S_\triangle}{a+c-b},$$

$$r_c = \frac{S_\triangle}{p-c} = \frac{2S_\triangle}{a+b-c}.$$

12.5. Соотношения между сторонами правильного и прямоугольного треугольников и радиусами вписанной и описанной окружностей. Сторона правильного треугольника a связана с радиусом описанной окружности R и радиусом вписанной окружности r соотношениями

$$R = \frac{a\sqrt{3}}{3},$$

$$r = \frac{a\sqrt{3}}{6}.$$

Центр вписанной в правильный треугольник окружности совпадает с центром описанной около него окружности и называется *центром* правильного треугольника.

Центр окружности, описанной вокруг прямоугольного треугольника, лежит на середине гипотенузы, и, следовательно, гипотенуза прямоугольного треугольника является диаметром окружности, описанной около прямоугольного треугольника.

Все точки окружности с диаметром AB (за исключением точек A и B) являются вершинами прямоугольных треугольников с гипотенузой AB (рис. 5.62).

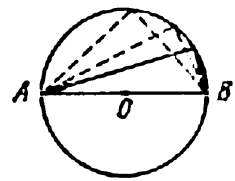


Рис. 5.62

12.6. Вписанные четырехугольники. Четырехугольник, все вершины которого принадлежат окружности, называется *вписанным* в эту окружность, а окружность — *описанной* вокруг этого четырехугольника.

Для того чтобы около выпуклого четырехугольника можно было описать окружность, необходимо и достаточно, чтобы сумма противолежащих углов этого четырехугольника была равна 180°.

В частности, из всех параллелограммов лишь около прямоугольника (квадрата) можно описать окружность.

Около трапеции можно описать окружность только тогда, когда она равнобочная.

Если четырехугольник $ABCD$ можно вписать в окружность, то произведение диагоналей этого четырехугольника равно сумме произведений его противоположных сторон (рис. 5.63):

$$|AC| \cdot |BD| = |AB| \cdot |CD| + |AD| \cdot |BC|.$$

12.7 Описанные четырехугольники. Четырехугольник, все стороны которого касаются окружности, называется *описанным* около окружности, а окружность — *вписанной* в этот четырехугольник.

Для того чтобы в выпуклый четырехугольник можно было вписать окружность, необходимо и достаточно, чтобы суммы противо-

лежащих сторон этого четырехугольника были равны. На рис. 5.64

$$|AB|+|CD|=|BC|+|AD|.$$

Из всех параллелограммов лишь в ромб (в частности, в квадрат)

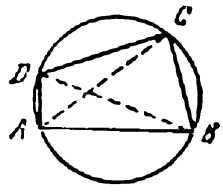


Рис. 5.63

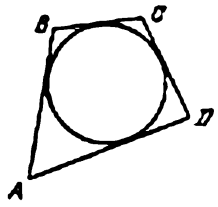


Рис. 5.64

можно вписать окружность. Центр вписанной окружности совпадает с точкой пересечения диагоналей указанных четырехугольников.

§ 13. Геометрические построения

Задачи на геометрические построения — один из традиционных разделов геометрии. Геометрические построения выполняются с помощью «односторонней» линейки и циркуля. Под «односторонней» линейкой понимается инструмент, с помощью которого можно выполнять единственную процедуру — проводить прямую (луч, отрезок) через две данные точки. Под циркулем понимается инструмент, которым можно строить окружности и откладывать на прямой геометрически заданный отрезок. В приведенных ниже задачах слова «отрезок задан» и «угол задан» означают, что нам дано геометрическое изображение отрезка (соответственно угла), а не их численное значение. Следует иметь в виду, что задача построения некоторой геометрической фигуры заключается не в практическом вычерчивании этой фигуры с известной степенью точности, а в том, как при помощи линейки и циркуля необходимое построение может быть выполнено теоретически в предположении, что наши инструменты дают абсолютную точность построения.

Сформулируем три классические задачи построения фигур, решение которых отыскивалось в течение нескольких столетий до тех пор, пока не было доказано, что эти задачи не могут быть решены с помощью линейки и циркуля.

1. Удвоение куба. Если данный куб имеет ребро, равное единице длины, то его объем будет равен кубической единице. Построить ребро куба, объем которого вдвое больше.

2. Трисекция угла. Разделить произвольный угол на три равных угла.

3. Квадратура круга. Построить квадрат (сторону квадрата), площадь которого равна площади круга, радиус которого принимается за единицу длины.

13.1. Построение прямых, параллельных и перпендикулярных данной прямой (данному отрезку).

1. *Построить прямую, параллельную данной прямой AB и проходящую через данную точку C.*

Произвольным раствором циркуля проводим окружность с центром в данной точке C так, чтобы она пересекла данную прямую AB

(рис. 5.65). Тем же раствором циркуля от одной из точек пересечения прямой и окружности (на рис. 5.65 точка M) откладываем на прямой AB в любую сторону отрезок MN. Снова тем же раствором засекаем из точки N точку окружности D. Через точки C и D проводим прямую. Это и есть искомая прямая.

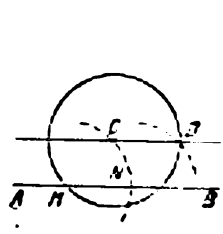


Рис. 5.65

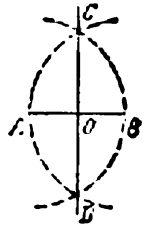


Рис. 5.66

2. *Разделить данный отрезок пополам и построить перпендикуляр к отрезку в его середине.*

Из концов данного отрезка AB как из центров (рис. 5.66) одним и тем же произвольным (бо́льшим $\frac{1}{2}|AB|$) радиусом строим две пересекающиеся дуги. Прямая, проходящая через точки пересечения дуг C и D,— искомый перпендикуляр.

3. *Построить перпендикуляр к данной прямой MN в данной точке A (построение прямого угла).*

Возьмем произвольную точку O, не принадлежащую данной прямой MN (рис. 5.67), и проведем окружность с центром в точке O радиуса OA. Через вторую точку пересечения окружности с прямой MN (на рис. 5.67 точка B) проводим диаметр BC. Прямая, проходящая через точки C и A, и есть искомый перпендикуляр. Угол, образованный лучами AC и AB,— прямой.

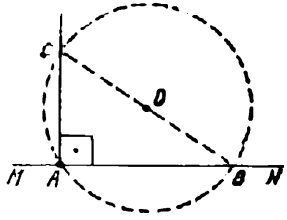


Рис. 5.67

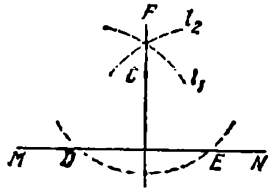


Рис. 5.68

4. *Опустить перпендикуляр из данной точки C на данную прямую MN.*

Из точки C как из центра (рис. 5.68) произвольным радиусом проводим дугу, пересекающую прямую MN в точках D и E. Из точек D и E как из центров проводим одним и тем же произвольным (но бо́льшим предыдущего) радиусом две дуги l_1, l_2, пересекающиеся в точке F. Прямая, проходящая через точки C и F, и есть искомый перпендикуляр.

13.2. Построение углов.

1. *Построить угол, равный данному углу MNK* (рис. 5.69, *а*).

Выбираем на плоскости произвольную точку O и проводим луч OA (рис. 5.69, *б*). Из вершины N данного угла как из центра описываем дугу PQ произвольного радиуса. Тем же раствором циркуля описываем из центра O дугу P_1Q_1 (точка Q_1 принадлежит лучу OA, а точка P_1 пока не фиксирована). Из точки Q_1 как из центра радиу-

сом, равным длине хорды PQ, засекаем точку P_1. Проводя луч OP_1, получаем угол P_1OQ_1, равный данному углу MNK.

2. *Построить углы, равные* 60° *и* 30°.

Из концов A и B произвольного отрезка AB как из центров радиусом, равным $|AB|$, описываем две дуги, пересекающиеся в точках C и D (рис. 5.70). Строим отрезки CD и AC. Обозначим точку пересечения отрезков AB и CD буквой O. Угол ACO равен 30°, а угол CAO равен 60°.

3. *Построить угол, равный* 45°,

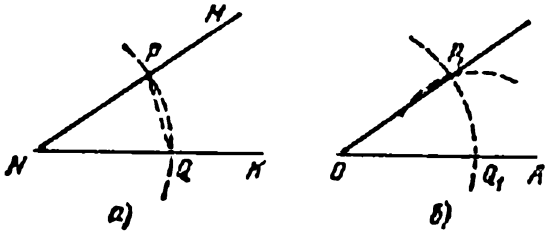


Рис. 5.69

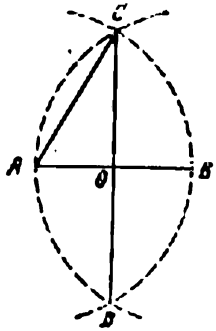


Рис. 5.70

На сторонах прямого угла AOB (рис. 5.71) откладываем равные отрезки OA и OB. Через точки A и B проводим прямую AB. Угол, образованный лучами BA и BO, равен 45°.

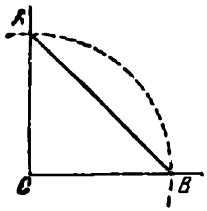


Рис. 5.71

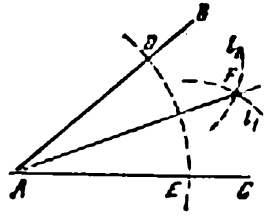


Рис. 5.72

4. *Разделить данный угол* BAC на два равных угла (построить биссектрису угла).

Из вершины угла BAC как из центра проводим дугу DE окружности произвольного радиуса (рис. 5.72). Из точек D и E пересечения этой дуги с лучами AB и AC описываем произвольным радиусом пересекающиеся дуги l_1 и l_2. Через точку их пересечения F проводим луч AF. Полученные углы BAF и FAC равны, а луч AF есть биссектриса угла BAC.

13.3. Построение отрезков.

1. *Разделить данный отрезок* AB *на заданное число равных отрезков.*

Строим прямую, параллельную прямой, содержащей данный отрезок AB (рис. 5.73). Берем произвольную точку A_1, принадлежащую построенной прямой. Из точки A_1 откладываем столько равных отрезков A_1C_1, C_1D_1, D_1E_1, ... произвольной длины, на сколько частей нужно разделить отрезок AB. Конец последнего отрезка обозначаем буквой B_1. Проводим прямые AA_1 и BB_1 до пересечения в точке O. Через пары точек $(C_1; O)$, $(D_1; O)$, $(E_1; O)$, ... проводим лучи, кото-

рые пересекут прямую AB в точках $C, D, E, \ldots$; эти точки и делят отрезок AB на нужное число равных отрезков.

2. *Разделить данный отрезок на отрезки, пропорциональные заданным величинам.*

Задача решается так же, как и предыдущая, только от точки A_1 откладываются отрезки, пропорциональные заданным величинам.

3. *Построить отрезок, средний пропорциональный (средний геометрический) двум данным отрезкам* *).

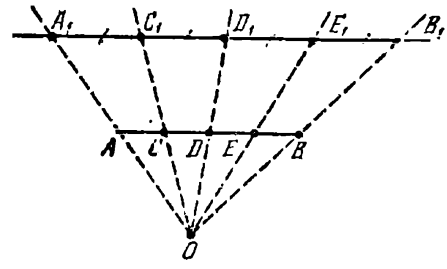


Рис. 5.73

На произвольной прямой отложим данные отрезки так, чтобы конец одного отрезка совпал с началом другого и эта точка была единственной их общей точкой (на рис. 5.74 отрезки AB и BC). Разделим отрезок AC пополам и радиусом, равным половине отрезка AC, построим окружность с центром в середине отрезка AC (на рис. 5.74 точка O — середина отрезка AC). Из точки B восставим перпендикуляр к отрезку AC. Точку пересечения перпендикуляра и окружности обозначим буквой D. Отрезок BD — средний пропорциональный отрезкам AB и BC:

$$\frac{|AB|}{|BD|}=\frac{|BD|}{|BC|}.$$

4. *Построить четвертый пропорциональный отрезок (по данным отрезкам a, b, c построить отрезок x такой, что $a:b=c:x$).*

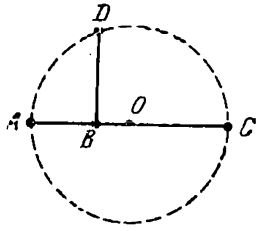


Рис. 5.74

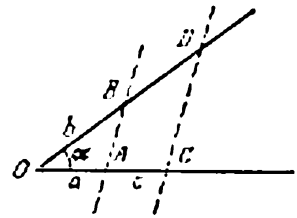


Рис. 5.75

На одной стороне произвольного угла α (рис. 5.75) от его вершины последовательно откладываем отрезки a и c, на другой — отрезок b. Через точку C проводим прямую, параллельную прямой AB. Отрезок BD, отсекаемый на стороне угла OB, — искомый.

5. *Построить отрезок, соизмеримый с данным* **).

---

*) Отрезок x, удовлетворяющий равенству $a/x=x/b$, где a и b — данные отрезки, называется отрезком, *средним пропорциональным (средним геометрическим)* двум данным отрезкам.

**) Отрезок AB называется *соизмеримым* с отрезком CD, если отношение длин этих отрезков — рациональное число:

$$\frac{|AB|}{|CD|}=\frac{m}{n} \qquad (m,\ n \text{ — натуральные числа}).$$

Пусть дан некоторый отрезок AB, и пусть требуется построить отрезок, соизмеримый с данным. Разделим отрезок AB на m равных частей. Возьмем один из этих отрезков и отложим его n раз на луче

Рис. 5.76

CN (рис. 5.76) так, чтобы конец предшествующего отрезка (за исключением первого) был началом следующего. Тогда конец последнего отрезка даст точку D — конец искомого отрезка CD.

6. Примеры построения отрезков, несоизмеримых с данным.

а) *По данному отрезку a построить отрезок $a\sqrt{2}$.*

Строим равнобедренный прямоугольный треугольник с катетами длиной a (рис. 5.77). Гипотенузой такого треугольника будет отрезок $a\sqrt{2}$.

б) *По данному отрезку a построить отрезок $a\sqrt{3}$.*

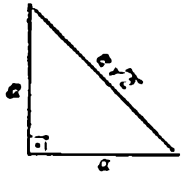


Рис. 5.77

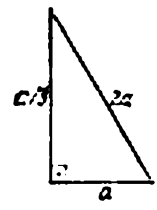


Рис. 5.78

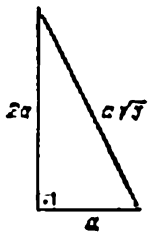


Рис. 5.79

Строим прямоугольный треугольник с катетом a и гипотенузой $2a$ (рис. 5.78).

Вторым катетом построенного прямоугольного треугольника будет отрезок $a\sqrt{3}$.

в) *По данному отрезку a построить отрезок $a\sqrt{5}$.*

Строим прямоугольный треугольник с катетами a и $2a$ (рис. 5.79). Гипотенузой построенного треугольника и будет отрезок $a\sqrt{5}$.

7. Примеры построения отрезков, задаваемых алгебраическими выражениями (под словами «отрезок задается алгебраическим выражением» понимается, что искомый отрезок задается как алгебраическая функция данных отрезков).

Для построения отрезков, задаваемых алгебраическими выражениями, используются следующие элементарные построения:

а) построение отрезка, среднего пропорционального двум данным отрезкам;

б) построение четвертого пропорционального отрезка;

в) построение отрезка, несоизмеримого с данным.

Пример 1. *Построить отрезок, задаваемый выражением*

$$\sqrt{a^2+ab+c^2},$$

где a, b, c — заданные отрезки.

Записывая выражение, стоящее под знаком радикала, в виде

$$a^2+ab+c^2=a(a+b)+c^2,$$

нетрудно заметить, что если мы построим отрезок x такой, что $x^2=a(a+b)$, т. е. отрезок, средний пропорциональный отрезкам a и $a+b$, то искомый отрезок будет получаться как гипотенуза прямоугольного треугольника с катетами x и c^2. Таким образом, построение искомого отрезка свелось к выполнению двух последовательных построений — построению среднего пропорционального отрезка и построению прямоугольного треугольника по двум катетам.

Пример 2. *Построить отрезок, задаваемый выражением*

$$\sqrt{3a^2+5b^2},$$

где a, b — заданные отрезки.

Строим отрезок $a\sqrt{3}$ как катет прямоугольного треугольника с заданными гипотенузой $2a$ и катетом a. Строим отрезок $b\sqrt{5}$ как гипотенузу прямоугольного треугольника с заданными катетами $2b$ и b. Искомый отрезок строится как гипотенуза прямоугольного треугольника с катетами $a\sqrt{3}$ и $b\sqrt{5}$.

Пример 3. *Построить отрезок x, задаваемый соотношением*

$$\frac{1}{x}=\frac{1}{a}+\frac{1}{b},$$

где a, b — заданные отрезки.

Проделав очевидные преобразования данного равенства:

$$\frac{1}{x}=\frac{1}{a}+\frac{1}{b}=\frac{a+b}{ab}\Leftrightarrow\frac{b}{x}=\frac{a+b}{a},$$

замечаем, что построение отрезка x сводится к построению четвертого пропорционального отрезка.

13.4. Построение окружностей и дуг окружностей.

1. *Через две данные точки провести окружность данного радиуса.*

Из данных точек A и B как из центров данным радиусом r проводим две пересекающиеся дуги l_1 и l_2. Точка их пересечения O (рис. 5.80) и есть центр искомой окружности *).

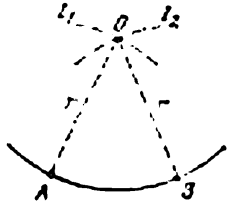


Рис. 5.80

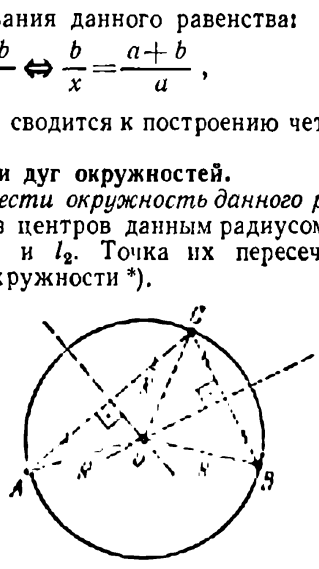


Рис. 5.81

2. *Через три данные точки A, B, C, не лежащие на одной прямой, провести окружность.*

Строим отрезки BC и AC (рис. 5.81), концами которых служат точки A, B, C, затем строим перпендикуляры к отрезкам BC и AC,

*) Здесь возможны три случая: существуют две окружности, проходящие через точки A и B (если $|AB|<2r$); если $|AB|=2r$ — такая окружность одна; при $|AB|>2r$ искомой окружности не существует.

проходящие через их середины. Точка пересечения этих перпендикуляров (на рис. 5.81 точка O) и будет центром искомой окружности. Радиус данной окружности равен расстоянию от точки O до любой из трех равноудаленных от O точек A, B, C.

3. *Через две данные точки A и B провести окружность, касающуюся данной прямой a, не проходящей через A и B.*

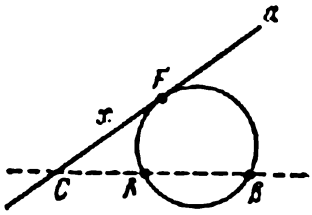


Рис. 5.82

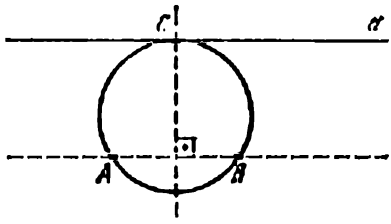


Рис. 5.83

Проводим прямую через данные точки A и B. Возможны два случая:

а) прямая AB пересекается с данной прямой*) a;

б) прямая AB параллельна данной прямой a.

Рассмотрим случай а). Обозначим точку пересечения прямых AB и a буквой C (рис. 5.82). Строим отрезок x, средний пропорциональный отрезкам CA и CB. Откладываем отрезок x на прямой a от точки C (на рис. 5.82 $x = CF$). Точка F — точка касания искомой окружности с данной прямой. Таким образом, решение данной задачи сводится к задаче построения окружности, проходящей через три точки A, B, F, не лежащие на одной прямой. Задача имеет и второе решение: можно построить вторую окружность, проходящую через данные точки A и B и касающуюся данной прямой a. Для этого отрезок a нужно отложить на прямой a по другую сторону от точки C.

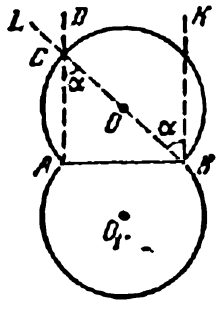


Рис. 5.84

Случай б). Если прямые AB и a параллельны, то проводим перпендикуляр через середину отрезка AB до пересечения с прямой a в точке C (рис. 5.83). Точка C — точка касания искомой окружности и данной прямой a. Окружность, проходящая через точки A, B, C — искомая.

4. *Найти центр данной дуги окружности.*

На данной дуге выбираем любые три точки A, B, C и строим центр окружности, проходящей через три заданные точки.

5. *Разделить данную дугу окружности на две равные части.*

Концы данной дуги соединяем хордой. Строим перпендикуляр к хорде, проходящий через ее середину. Точка пересечения перпендикуляра и дуги делит данную дугу на две равные дуги.

6. *Найти множество точек плоскости, из которых данный отрезок AB виден под данным острым углом α.*

---

*) Считаем, что точки A и B лежат по одну сторону от прямой a. Если точки A и B лежат по разные стороны прямой, то задача не имеет решений.

Множеством точек плоскости, из которых данный отрезок AB виден под данным углом α, является дуга окружности величиной $2\pi - 2\alpha$, стягиваемая хордой AB. Центр и радиус этой окружности находятся следующим образом. Из точек A и B (рис. 5.84) восставим перпендикуляры к отрезку AB. Обозначим эти перпендикуляры AD и BK соответственно. От луча BK отложим угол $KBL = \alpha$ так, чтобы луч BL пересек луч AD в некоторой точке C. Середина O отрезка BC — центр искомой окружности, $\frac{1}{2}|BC|$ — ее радиус. Из любой точки, лежащей на дуге ABC, отрезок AB виден под углом α, так как любой угол, опирающийся на хорду AB, с вершиной на дуге ACB, измеряется половиной дуги, на которую он опирается.

Задача допускает и второе решение — существует вторая окружность такого же радиуса (на рис. 5.84 — лежащая ниже отрезка AB), из которой данный отрезок AB виден под углом α.

13.5. Построение касательных к окружностям.

1. *Провести через данную точку касательную к данной окружности.*

Если данная точка A принадлежит данной окружности с центром O (рис. 5.85), то искомой касательной будет перпендикуляр к AO, проведенной через конец A отрезка AO.

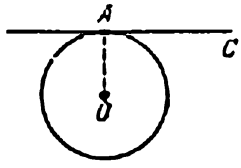


Рис. 5.85

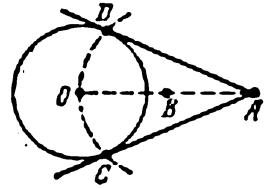


Рис. 5.86

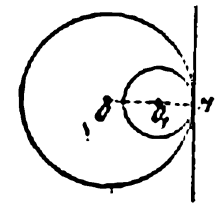


Рис. 5.87

Если точка A лежит вне круга (рис. 5.86), то необходимо разделить отрезок AO пополам и из его середины B радиусом OB провести дугу DOC (D и C — точки пересечения дуги и данной окружности). Через пары точек $(D; A)$ и $(C; A)$ проводим прямые, которые будут искомыми касательными.

2. *Провести общую внешнюю касательную к данным двум окружностям.*

Задача не имеет решений, если один круг целиком содержится в другом и окружности не касаются. Если один круг целиком содержится в другом и окружности касаются, то задача имеет единственное решение (существует только одна прямая, касающаяся обеих окружностей). Во всех остальных случаях взаимного расположения двух окружностей задача имеет два решения (т. е. существуют две различные общие внешние касательные к данным окружностям).

Рассмотрим случай внутреннего касания двух окружностей. Проводим линию центров OO_1 этих окружностей (рис. 5.87). В точке касания двух окружностей (на рис. 5.87 точка M) строим прямую, перпендикулярную линии центров. Это и есть искомая касательная.

Рассмотрим случаи, в которых задача имеет два решения.

а) Если радиусы данных окружностей равны, то проведем через центры окружностей O и O_1 (рис. 5.88) диаметры AB и A_1B_1, перпендикулярные линии центров OO_1. Прямые, проходящие через пары точек $(A; A_1)$ и $(B; B_1)$, — искомые касательные.

б) Если радиусы R и R_1 данных окружностей не равны ($R > R_1$), то строим окружность радиуса $OC = R - R_1$ с центром, совпадающим с центром большего круга (рис. 5.89). К построенной окружности проводим касательные AC и AC_1, проходящие через центр меньшего круга A. Через центр большей окружности O и точки касания C и C_1 проводим лучи OC и OC_1, которые пересекут большую окружность в точках D и D_1. Проводим радиусы AE и AE_1, перпендикулярные прямым AC и AC_1 соответственно. Прямые, проходящие через пары точек $(D; E)$, $(D_1; E_1)$, — искомые касательные.

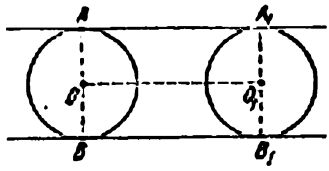


Рис. 5.88

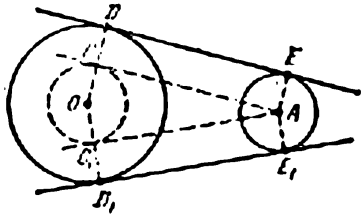


Рис. 5.89

3. *Провести общую внутреннюю касательную к двум данным окружностям.*

Задача не имеет решения, если окружности пересекаются или если один круг целиком содержится в другом. Если окружности касаются друг друга извне, задача имеет единственное решение — через точку касания проводится перпендикуляр к линии центров (рис. 5.90).

Если окружности не пересекаются, то задача имеет два решения (т. е. существуют две различные прямые, каждая из которых будет касательной как к одной, так и к другой окружности). Касательные

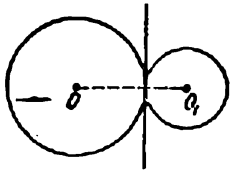


Рис. 5.90

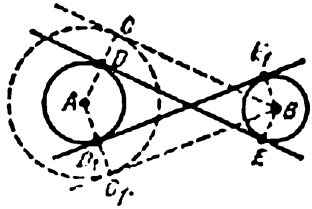


Рис. 5.91

строятся так. Строим окружность радиусом, равным сумме радиусов данных окружностей, и с центром, совпадающим с центром одной из окружностей (рис. 5.91). Из центра B второй окружности проводим касательные BC и BC_1 к построенной окружности. Проводим радиусы AC и AC_1 в точки касания. Эти радиусы пересекут окружность с центром A в точках D и D_1. Через центр второй окружности B проводим радиусы BE и BE_1, перпендикулярные прямым BC и BC_1 соответственно. Прямые, проходящие через пары точек $(D; E)$ и $(D_1; E_1)$, — искомые касательные.

13.6. Построение окружности, описанной около многоугольника, и многоугольника, вписанного в окружность.

1. *Описать окружность около данного треугольника.*

Через вершины треугольника A, B, C проводится окружность,

2. *Описать окружность около данного прямоугольника (или квадрата).*

Проводим диагонали AC и BD (рис. 5.92). Из точки O их пересечения как из центра проводим окружность радиусом $|OA|$. Построенная окружность — искомая.

3. *Вписать квадрат в данный круг.*

Проводим два взаимно перпендикулярных диаметра AB и CD. Четырехугольник с вершинами A, B, C, D — искомый квадрат (рис. 5.93).

4. *Вписать в данную окружность правильный шестиугольник и правильный треугольник.*

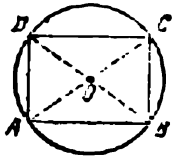


Рис. 5.92

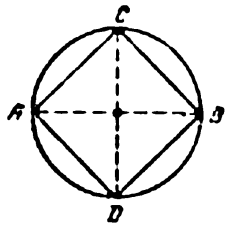


Рис. 5.93

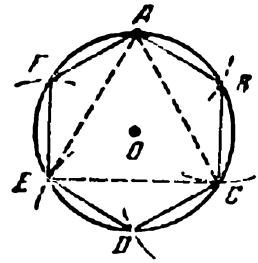


Рис. 5.94

Выберем на окружности произвольную точку (скажем, точку A) и раствором циркуля, равным радиусу окружности, делаем на окружности засечки, получая последовательно точки B, C, D, E, F (рис. 5.94). Последовательно соединяя указанные точки, получаем правильный шестиугольник, вписанный в данную окружность. Соединяя их через одну, получим правильный (равносторонний) треугольник, вписанный в данную окружность.

5. *Вписать в данную окружность правильный восьмиугольник.*

Проводим два взаимно перпендикулярных диаметра AB и CD (рис. 5.95). Проведя биссектрисы четырех прямых углов, получим точки пересечения биссектрис с окружностью (на рис. 5.95 — точки E, F, M, N). Последовательно соединяя полученные восемь точек A, E, C, F, B, M, D, N, получаем искомый восьмиугольник.

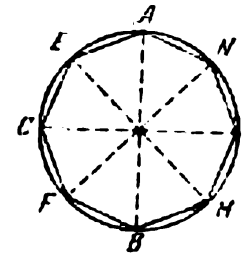


Рис. 5.95

13.7. Построение окружности, вписанной в многоугольник, и многоугольника, описанного около окружности.

1. *Вписать окружность в данный треугольник.*

Строим биссектрисы двух внутренних углов треугольника. Из точки O пересечения биссектрис (рис. 5.96) проводим перпендикуляр к любой стороне треугольника (например, к стороне AB), который пересечет эту сторону в точке D. Радиусом, равным $|OD|$, из точки O как из центра описываем искомую окружность.

2. *Вписать окружность в ромб (или квадрат).*

Проводим диагонали ромба (на рис. 5.97 — отрезки AC и BD). Через точку O их пересечения проводим перпендикуляр к любой

стороне ромба (на рис. 5.97 — к стороне BC); E — точка пересечения перпендикуляра со стороной. Окружность с центром O и радиусом, равным $|OE|$, — искомая.

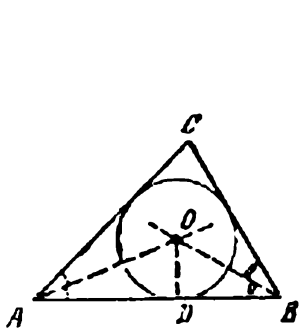


Рис. 5.96

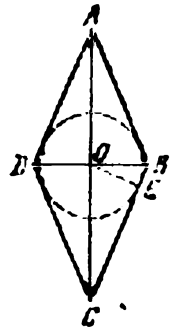


Рис. 5.97

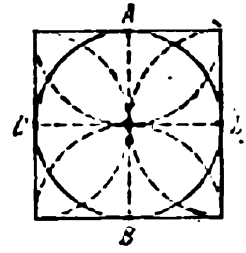


Рис. 5.98

3. *Описать квадрат около данной окружности.*

Проводим два взаимно перпендикулярных диаметра (на рис. 5.98 AB и CD — диаметры). Из их концов (точек A, B, C, D) как из центров описываем четыре полуокружности радиусами, равными радиусу данной окружности. Точки пересечения полуокружностей — вершины искомого квадрата.

4. *Описать вокруг данной окружности правильный шестиугольник и правильный восьмиугольник.*

Отмечая на окружности соответственно шесть (восемь) точек так, как это сделано в задачах 4, 5 п. 13.6, строим касательные к данной окружности в отмеченных точках. Точки пересечения соседних касательных дадут вершины правильного шестиугольника (восьмиугольника).

13.8. Построение треугольников.

1. *Построить треугольник по трем сторонам a, b, c.*

Из концов отрезка AB ($|AB| = a$) как из центров проводим две дуги радиусов b и c (рис. 5.99). Точка их пересечения C — третья

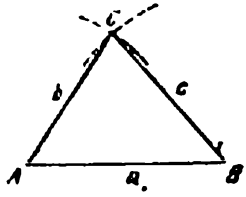


Рис. 5.99

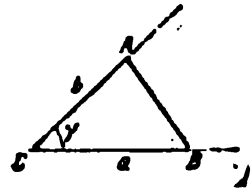


Рис. 5.100

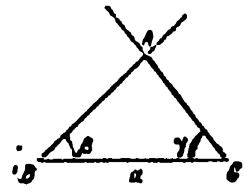


Рис. 5.101

вершина искомого треугольника ABC. Задача имеет решение, если отрезки a, b, c удовлетворяют неравенствам треугольника.

2. *Построить треугольник по двум сторонам a, b и углу γ между ними.*

Строим угол, равный углу γ. От его вершины C на сторонах угла откладываем отрезки CA и CB, соответственно равные отрезкам b и a (рис. 5.100). Соединяем точки A и B. Треугольник ABC — искомый.

3. *Построить треугольник по стороне a и прилежащим к ней углам* β *и* γ ($\beta+\gamma<180°$).

При каждом из концов отрезка BC ($|BC|=a$) откладываем углы β и γ (рис. 5.101). Точка A пересечения лучей BA и CA есть третья вершина искомого треугольника ABC.

4. *Построить треугольник по основанию a, высоте* h_a *и углу* α *при вершине.*

Строим дугу окружности, из которой отрезок AB ($|AB|=a$) виден под данным углом α (см. задачу 6 п. 13.4). Строим произволь-

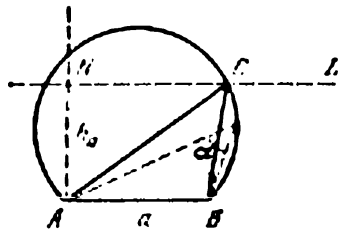


Рис. 5.102

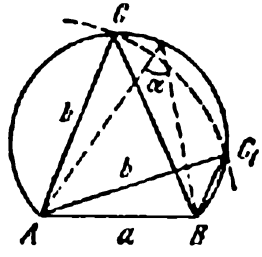


Рис. 5.103

ный перпендикуляр к отрезку AB и откладываем на нем отрезок AN ($|AN|=h_a$). Через точку N проводим прямую NL, параллельную прямой, проходящей через точки A, B. Полученная точка пересечения прямой NL с дугой окружности (точка C на рис. 5.102) — вершина искомого треугольника ABC (задача имеет два решения).

5. *Построить треугольник по двум сторонам a, b и углу* α, *противолежащему стороне a.*

Строим дугу окружности, из которой отрезок AB ($|AB|=a$) виден под данным углом α. Из конца A отрезка AB как из центра радиусом, равным b, проводим дугу, пересекающуюся с первой дугой. Точка пересечения этих дуг (на рис. 5.103 — точки C и C_1) и будет третьей вершиной искомого треугольника ABC (C_1). В зависимости от величин a, b и α задача имеет одно или два решения или вовсе не имеет решений.

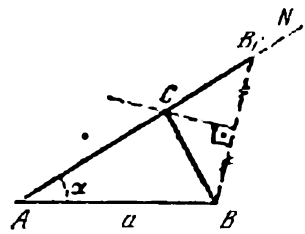


Рис. 5.104

6. *Построить треугольник по стороне a, углу* α *при основании и сумме двух других сторон* $b+c$.

От конца A отрезка AB ($|AB|=a$) откладываем угол α. На луче AN откладываем данную сумму двух сторон искомого треугольника (на рис. 5.104 — отрезок AB_1 ($|AB_1|=b+c$)). Из середины отрезка BB_1 восставляем перпендикуляр до пересечения с лучом AB_1. Точка пересечения луча и перпендикуляра (на рис. 5.104 — точка C) — третья вершина искомого треугольника ABC.

7. *Построить треугольник по стороне a, углу при основании* α *и разности двух других сторон* $b-c$.

От конца A отрезка AB ($|AB|=a$) откладываем угол α. На луче AN откладываем данную разность $b-c$ (на рис. 5.105 — отрезок AB_1). Из середины отрезка BB_1 восставляем перпендикуляр до пересечения с лучом AN. Точка пересечения луча AN и перпендикуляра

(на рис. 5.105 — точка C) — третья вершина искомого треугольника ABC.

8. *Построить треугольник по двум данным углам* α, β *и периметру* P.

Строим углы $\alpha/2$, $\beta/2$. Строим отрезок AB ($|AB| = P$). От концов A и B откладываем углы $\alpha/2$ и $\beta/2$ так, чтобы лучи, являющиеся вторыми сторонами этих углов, принадлежали одной полуплоскости относительно прямой AB (рис. 5.106). Из середин отрезков AE и BE восставляем перпендикуляры до пересечения с отрезком AB в точках C и D. Треугольник CED — искомый.

9. *Построить треугольник по высоте* h_a, *медиане* m_a *и биссектрисе* l_a, *проведенным из одной вершины* *).

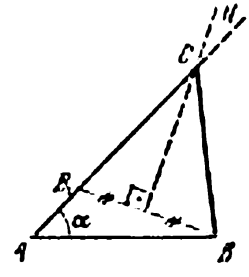


Рис. 5.105

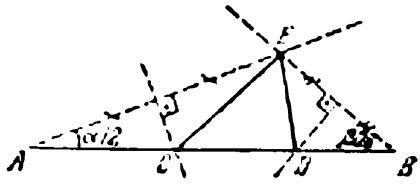


Рис. 5.106

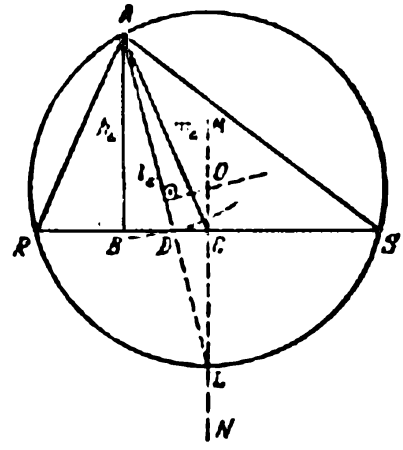


Рис. 5.107

Строим прямоугольный треугольник ABC с гипотенузой AC ($|AC| = m_a$) и катетом AB ($|AB| = h_a$) (рис. 5.107). Проводим из центра A дугу радиусом l_a, пересекающую катет BC в точке D. Строим прямую MN, перпендикулярную прямой BC и проходящую через точку C. Обозначаем точку пересечения луча AD и прямой MN буквой L. Восставляем перпендикуляр к отрезку AL в его середине. Точка пересечения O данного перпендикуляра и прямой MN — центр окружности, описанной вокруг искомого треугольника. Радиусом AO из центра O проводим окружность, пересекающую прямую BC в точках R и S. Треугольник с вершинами A, R, S — искомый.

10. *Построить треугольник по трем медианам* m_a, m_b, m_c.

Строим отрезки $\frac{2}{3}m_a$, $\frac{2}{3}m_b$, $\frac{2}{3}m_c$. Строим треугольник APQ (рис. 5.108) со сторонами $\frac{2}{3}m_a$, $\frac{2}{3}m_b$, $\frac{2}{3}m_c$. На луче AP от точки P откладываем отрезок $PM = \frac{1}{3}m_a$. Находим середину N отрезка PQ и на луче AN откладываем отрезок $NC = AN$. Строим луч CM

*) Задача имеет единственное решение, если $h_a < l_a < m_a$.

и откладываем на нем отрезок $BM = CM$ так, чтобы точка M принадлежала отрезку BC. Точки A, B, C — вершины искомого треугольника.

11. *Построение прямоугольных треугольников.*

Построение прямоугольного треугольника по двум катетам проводится так же, как в задаче 2 настоящего пункта для случая $\gamma = 90°$.

Построение прямоугольного треугольника по катету и острому углу β проводится так же, как в задаче 3, если угол γ положить равным 90°.

Построение прямоугольного треугольника по гипотенузе и острому углу α сводится к задаче 3 для случая острых углов α и $\beta = 90° - \alpha$.

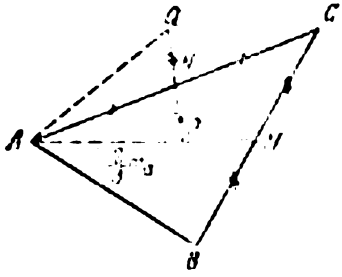


Рис. 5.108

12. *Построить параллелограмм по данным сторонам a, b и углу α между ними.*

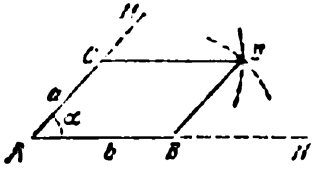


Рис. 5.109

Строим угол MAN, равный данному углу α (рис. 5.109). На луче AM откладываем отрезок $AC = a$, а на луче AN — отрезок AB b. Проводим из точки B как из центра дугу радиуса AC, а из точки C — дугу радиуса AB. Точку пересечения этих дуг (на рис. 5.109 точку D) соединяем с точками B и C. Четырехугольник $ABDC$ — искомый параллелограмм.

§ 14. Многогранный угол

Пусть даны плоский многоугольник Ф и точка S, не принадлежащая плоскости данного многоугольника (на рис. 5.110 многоугольник Ф — шестиугольник $ABCDEF$). Объединение всех лучей, имеющих общее начало S и пересекающих данный многоугольник Ф, называется *многогранным углом* *).

Рис. 5.110

Точка S называется *вершиной* многогранного угла; лучи SA, SB, SC, SD, SE, SF, содержащие вершины многоугольника, — его *ребрами*; плоскости, содержащие треугольники SAB, SBC и т. д., — его *гранями*.

В зависимости от числа граней различают трехгранные, четырехгранные, пятигранные и т. д. углы. Многогранный угол обозначают, отмечая сначала вершину, а затем последовательно по одной точке на каждом из ребер. Так, шестигранный угол, изображенный на рис. 5.110, обоз-

*) Иногда под *многогранным углом* понимают множество всех лучей, имеющих общее начало S и пересекающих не весь многоугольник, а лишь замкнутую ломаную $ABCDEF$.

начают $SABCDEF$. Каждые две грани многогранного угла, имеющие общее ребро, образуют двугранный угол.

Множество всех точек многогранного угла, не принадлежащих граням, называют его *внутренней областью*. Многогранный угол, внутренняя область которого расположена по одну сторону от плоскости каждой из его граней, называется *выпуклым* многогранным углом. В противном случае многогранный угол называется *невыпуклым*.

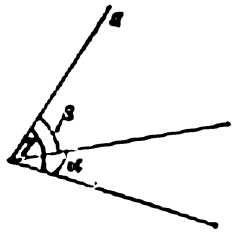


Рис. 5.111

Углы ASB, BSC и т. д. называются *плоскими углами* многогранного угла $SABCDEF$ (см. рис. 5.110).

Свойства плоских углов многогранного угла:

1) каждый плоский угол многогранного угла меньше суммы остальных его плоских углов;

2) в выпуклом многогранном угле сумма плоских углов меньше 360°.

Простейшим многогранным углом является трехгранный угол. Необходимым и достаточным условием существования трехгранного угла с данными плоскими углами α, β, γ является выполнение неравенств

$$|\beta-\gamma|<\alpha<\beta+\gamma, \quad \alpha+\beta+\gamma<360^\circ.$$

Теорема косинусов для трехгранного угла. Косинус плоского угла трехгранного угла равен произведению косинусов двух остальных плоских углов, сложенному с произведением синусов тех же углов и косинуса двугранного угла, определяемого этими плоскими углами (рис. 5.111):

$$\cos\alpha=\cos\beta\cos\gamma+\sin\beta\sin\gamma\cos a.$$

§ 15. Многогранная поверхность. Многогранник

Многогранной поверхностью называют такое объединение конечного числа плоских многоугольников, что каждая сторона любого из многоугольников является в то же время стороной другого (но только одного) многоугольника, называемого *смежным* с первым многоугольником.

От любого из многоугольников, составляющих многогранную поверхность, можно дойти до любого другого, двигаясь по смежным многоугольникам.

Многоугольники, составляющие многогранную поверхность, называются ее *гранями*; стороны многоугольников называются *ребрами*, а вершины — *вершинами* многогранной поверхности.

На рис. 5.112 изображены объединения многоугольников, удовлетворяющие указанным требованиям и являющиеся многогранными поверхностями. На рис. 5.113 изображены фигуры, не являющиеся многогранными поверхностями.

Многогранная поверхность делит пространство на две части — внутреннюю и внешнюю области многогранной поверхности. Из двух областей внешней будет та, в которой можно провести прямые, целиком принадлежащие области.

Объединение многогранной поверхности и ее внутренней области называют *многогранником*. При этом многогранную поверхность и ее

внутреннюю область называют соответственно *поверхностью* и *внутренней областью* многогранника. Грани, ребра и вершины поверхности многогранника называют соответственно *гранями*, *ребрами* и *вершинами* многогранника.

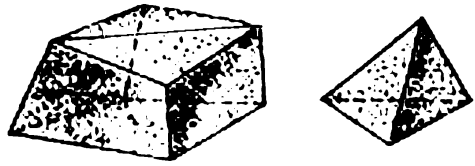


Рис. 5.112

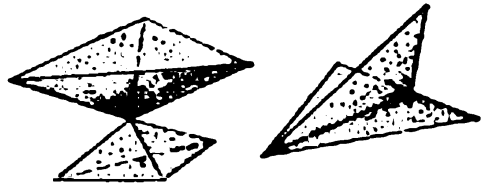


Рис. 5.113

Многогранник называется *выпуклым*, если он весь лежит по одну сторону от плоскости любой его грани.

Отрезок, соединяющий две вершины многогранника, не принадлежащие одной грани, называются *диагональю* многогранника.

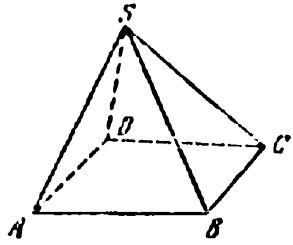


Рис. 5.114

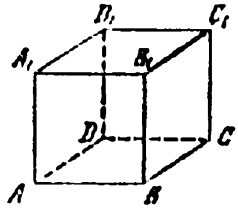


Рис. 5.115

Многогранник обычно обозначается перечислением его вершин и указанием его специальных свойств. Например, многогранник $SABCD$, изображенный на рис. 5.114, — пирамида; многогранник $ABCDA_1B_1C_1D_1$ (рис. 5.115) — параллелепипед.

§ 16. Призма

Многогранник, две грани которого — равные n-угольники, лежащие в параллельных плоскостях, а остальные n граней — параллелограммы, называется n-угольной *призмой*.

Пару равных n-угольников называют *основаниями* призмы. Остальные грани призмы называют ее *боковыми гранями*, а их объединение — *боковой поверхностью* призмы. На рис. 5.116 изображена пятиугольная призма.

Стороны граней призмы называют *ребрами*, а концы ребер — *вершинами* призмы. Ребра, не принадлежащие основанию призмы, называют *боковыми ребрами*.

Призму, боковые ребра которой перпендикулярны плоскостям оснований, называют *прямой* призмой. В противном случае призма называется *наклонной*.

Отрезок перпендикуляра к плоскостям оснований призмы, концы которого принадлежат этим плоскостям, называют *высотой* призмы.

Прямая призма, основанием которой является правильный многоугольник, называется *правильной* призмой.

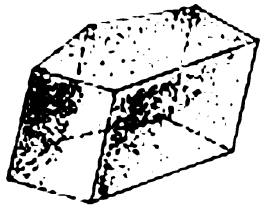


Рис. 5.116

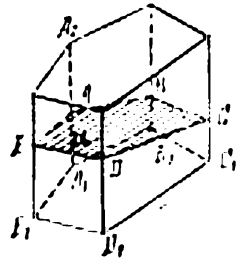


Рис. 5.117

Площадь боковой поверхности призмы. Пусть дана произвольная призма (на рис. 5.117 — пятиугольная призма). Через точку A, принадлежащую одному из ее боковых ребер, проведем плоскость α, перпендикулярную этому ребру (и, следовательно, перпендикулярную всем остальным боковым ребрам). Если плоскость α пересекает все боковые ребра призмы, то многоугольник, полученный в результате сечения всех боковых граней плоскостью α (на рис. 5.117 — пятиугольник $ABCDE$), называется *перпендикулярным сечением* призмы (если такого многоугольника не существует, то за перпендикулярное сечение призмы принимают многоугольник с вершинами в точках пересечения плоскости α с продолжениями боковых ребер).

Площадь боковой поверхности призмы вычисляется по формуле

$$S_{\text{бок}} = P_{\text{п}} \cdot |A_1A_2|,$$

где $P_{\text{п}}$ — периметр перпендикулярного сечения призмы, $|A_1A_2|$ — длина бокового ребра.

Площадь полной поверхности призмы равна сумме площадей боковой поверхности призмы и двух ее оснований.

Объем призмы вычисляется по формуле

$$V = S_{\text{п}} \cdot |A_1A_2|,$$

где $S_{\text{п}}$ — площадь перпендикулярного сечения призмы, $|A_1A_2|$ — длина бокового ребра, или по формуле

$$V = S_{\text{осн}} \cdot H,$$

где $S_{\text{осн}}$ — площадь основания призмы, H — высота.

§ 17. Параллелепипед. Куб

Параллелепипедом называется призма, основанием которой служат параллелограммы. Все шесть граней параллелепипеда (рис. 5.118) — параллелограммы. Отрезки, соединяющие вершины парал-

лепипеда, не принадлежащие одной и той же грани, называются *диагоналями* параллелепипеда.

Свойства параллелепипеда:

1) середина диагонали параллелепипеда является его центром симметрии;

2) противолежащие грани параллелепипеда попарно равны и параллельны;

3) все четыре диагонали параллелепипеда пересекаются в одной точке и делятся ею пополам.

Параллелепипед, боковые ребра которого перпендикулярны плоскости основания параллелепипеда, называется *прямым* параллелепипедом (на рис. 5.119 $ABCDA_1B_1C_1D_1$ — прямой параллелепипед).

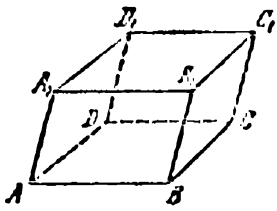


Рис. 5.118

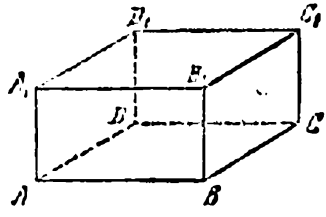


Рис. 5.119

Прямой параллелепипед, основанием которого служит прямоугольник, называется *прямоугольным* параллелепипедом. Все грани прямоугольного параллелепипеда — прямоугольники. Длины трех ребер прямоугольного параллелепипеда, выходящих из одной вершины, называются *измерениями* прямоугольного параллелепипеда.

Свойства прямоугольного параллелепипеда:

1) квадрат диагонали прямоугольного параллелепипеда равен сумме квадратов трех его измерений:

$$d^2 = a^2 + b^2 + c^2;$$

2) все диагонали прямоугольного параллелепипеда равны.

Так как параллелепипед есть частный случай призмы, то площадь поверхности и объем параллелепипеда вычисляются по формулам для площади поверхности и объема призмы. Кроме того, объем прямоугольного параллелепипеда можно вычислить по формуле

$$V = abc,$$

где a, b, c — три измерения прямоугольного параллелепипеда.

Прямоугольный параллелепипед с равными измерениями называется *кубом*. Все грани куба — равные квадраты.

Объем куба вычисляется по формуле

$$V = a^3,$$

где a — измерение куба.

§ 18. Пирамида. Усеченная пирамида

Многогранник, одна из граней которого — произвольный многоугольник, а остальные грани — треугольники, имеющие одну общую вершину, называется *пирамидой*. При этом многоугольник называется

основанием пирамиды, а остальные грани (треугольники) называются *боковыми гранями* пирамиды.

Различают треугольные, четырехугольные, пятиугольные и т. д. пирамиды в зависимости от вида многоугольника, лежащего в ее основании. На рис. 5.120 изображена четырехугольная пирамида $SABCD$ с основанием $ABCD$ и боковыми гранями SAB, SBC, SCD, SAD. Треугольную пирамиду также называют тетраэдром.

Стороны граней пирамиды называются *ребрами* пирамиды. Ребра, принадлежащие основанию пирамиды, называют *ребрами основания*, а все остальные ребра — *боковыми ребрами*. Общая вершина всех треугольников (боковых граней) называется *вершиной* пирамиды (на

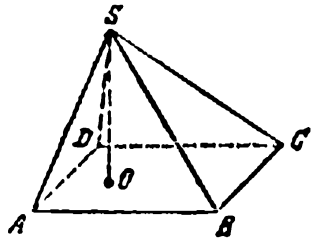


Рис. 5.120

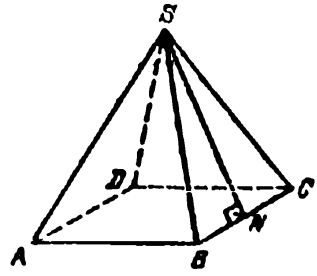


Рис. 5.121

рис. 5.120 точка S — вершина пирамиды, отрезки SA, SB, SC, SD — боковые ребра, отрезки AB, BC, CD, AD — ребра основания).

Высотой пирамиды называется отрезок перпендикуляра, проведенного из вершины пирамиды S к плоскости основания (концами этого отрезка являются вершина пирамиды и основание перпендикуляра). На рис. 5.120 SO — высота пирамиды.

Правильная пирамида. Пирамида называется *правильной*, если основанием пирамиды является правильный многоугольник, а ортогональная проекция вершины на плоскость основания совпадает с центром многоугольника, лежащего в основании пирамиды.

Свойства правильной пирамиды:

1) все боковые ребра правильной пирамиды равны между собой;

2) все боковые грани — равные равнобедренные треугольники.

Высота боковой грани правильной пирамиды, проведенная из ее вершины, называется *апофемой* этой пирамиды. На рис. 5.121 SN — апофема. Все апофемы правильной пирамиды равны между собой.

Площадь боковой поверхности пирамиды равна сумме площадей треугольников боковых граней, а площадь полной поверхности равна сумме площади боковой поверхности и площади основания.

Площадь боковой поверхности правильной пирамиды вычисляется по формуле

$$S = \frac{1}{2} Ph,$$

где P — периметр основания пирамиды, h — апофема.

Объем пирамиды вычисляется по формуле

$$V = \frac{1}{3} S_{\text{осн}} \cdot H,$$

где $S_{\text{осн}}$ — площадь основания пирамиды, H — высота пирамиды.

Усеченная пирамида. Возьмем произвольную пирамиду. Через точку, принадлежащую какому-нибудь боковому ребру, и не совпадающую с его концами, проведем плоскость, параллельную плоскости основания (рис. 5.122). Проведенная плоскость отсечет от данной пирамиды $SABCD$ (рис. 5.122) пирамиду $SA_1B_1C_1D_1$.

Многогранник, вершинами которого служат вершины основания данной пирамиды и вершины основания отсекаемой пирамиды, называется *усеченной пирамидой*. Основания усеченной пирамиды — гомотетичные многоугольники (на рис. 5.122 четырехугольники $ABCD$ и $A_1B_1C_1D_1$ — основания усеченной пирамиды). Центр гомотетии — вершина пирамиды. Отрезок перпендикуляра к плоскостям оснований, с концами на плоскостях оснований пирамиды, называется *высотой* усеченной пирамиды. Боковые грани усеченной пирамиды — трапеции.

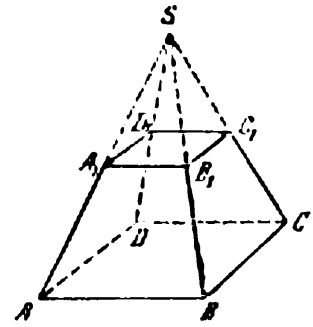


Рис. 5.122

Усеченная пирамида называется *правильной*, если она является частью правильной пирамиды. Боковые грани правильной усеченной пирамиды — равные равнобедренные трапеции. Высота каждой из этих трапеций называется *апофемой* правильной усеченной пирамиды.

Площадь боковой поверхности правильной усеченной пирамиды вычисляется по формуле

$$S=\frac{1}{2}(P+p)\,h,$$

где P, p — периметры оснований пирамиды, h — апофема.

Площадь полной поверхности усеченной пирамиды равна сумме площадей оснований и площади боковой поверхности пирамиды.

Объем усеченной пирамиды вычисляется по формуле

$$V=\frac{1}{3}H\,(S_1+\sqrt{S_1S_2}+S_2),$$

где H — высота усеченной пирамиды, S_1 и S_2 — площади оснований усеченной пирамиды.

§ 19. Правильные многогранники

Многогранник называется *правильным*, если все его грани — равные правильные многоугольники, а все многогранные углы имеют одинаковое число граней. Все ребра правильного многогранника — равные отрезки, все плоские углы правильного многогранника также равны.

Существует пять видов различных правильных многогранников (выпуклых): *куб*, правильный *тетраэдр*, правильный восьмигранник (правильный *октаэдр*), правильный двенадцатигранник (*додекаэдр*), правильный двадцатигранник (*икосаэдр*). На рис. 5.123—5.127 изображены перечисленные правильные многогранники, а также развертки их поверхностей.

Далее будем пользоваться следующими обозначениями: объем многогранника будем обозначать V; площадь поверхности — S; радиус описанной сферы — R; радиус вписанной сферы — r; высоту (для тех многогранников, для которых это понятие имеет смысл) — H; каждое из равных ребер — a.

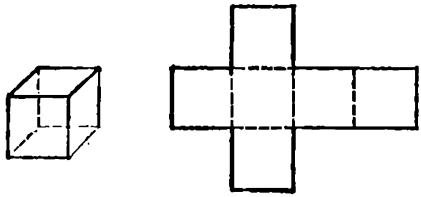


Рис. 5.123

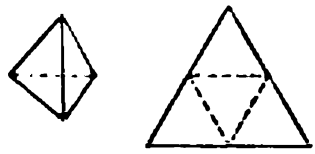


Рис. 5.124

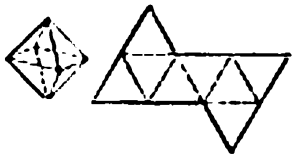


Рис. 5.125

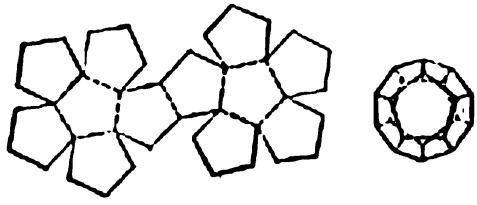


Рис. 5.126

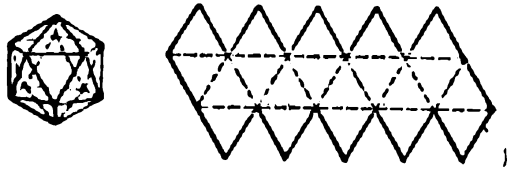


Рис. 5.127

Куб. Все шесть его граней — равные квадраты. Куб имеет восемь вершин и двенадцать ребер.

$$V=a^3, \qquad S=6a^2,$$

$$R=\frac{a\sqrt{3}}{2}, \qquad r=\frac{a}{2}.$$

Тетраэдр. Все четыре его грани — равносторонние равные треугольники. Тетраэдр имеет четыре вершины и шесть ребер.

$$V=\frac{a^3\sqrt{2}}{12}, \qquad S=a^2\sqrt{3},$$

$$R=\frac{a\sqrt{6}}{4}, \qquad r=\frac{a\sqrt{6}}{12}, \qquad H=\frac{a\sqrt{6}}{3}.$$

Октаэдр. Все восемь его граней — равносторонние равные треугольники. Октаэдр имеет шесть вершин и двенадцать ребер.

$$V=\frac{a^3\sqrt{2}}{3}, \qquad S=2a^2\sqrt{3}, \qquad R=\frac{a\sqrt{2}}{2}, \qquad r=\frac{a\sqrt{6}}{6}.$$

Додекаэдр. Все двенадцать его граней — правильные равные пятиугольники. Додекаэдр имеет двадцать вершин и тридцать ребер.

$$V=\frac{a^3(15+7\sqrt{5})}{4},\qquad S=3a^2\sqrt{5(5+2\sqrt{5})},$$

$$R=\frac{a\sqrt{3}(1+\sqrt{5})}{4},\qquad r=\frac{a\sqrt{10(25+11\sqrt{5})}}{20}.$$

Икосаэдр. Все двадцать его граней — равносторонние равные треугольники. Икосаэдр имеет двенадцать вершин и тридцать ребер.

$$V=\frac{5a^3(3+\sqrt{5})}{12},\qquad S=5a^2\sqrt{3},$$

$$R=\frac{a\sqrt{2(5+\sqrt{5})}}{4},\qquad r=\frac{a\sqrt{3}(3+\sqrt{5})}{12}.$$

Число ребер, число вершин и число граней правильных многогранников связаны формулой Эйлера *)

$$N-L+F=2,$$

где N — число вершин, L — число ребер, F — число граней.

§ 20. Цилиндр

Цилиндрической поверхностью называется поверхность, образованная прямой (*образующей*), перемещающейся параллельно заданному направлению в пространстве вдоль некоторой заданной линии (*направляющей*).

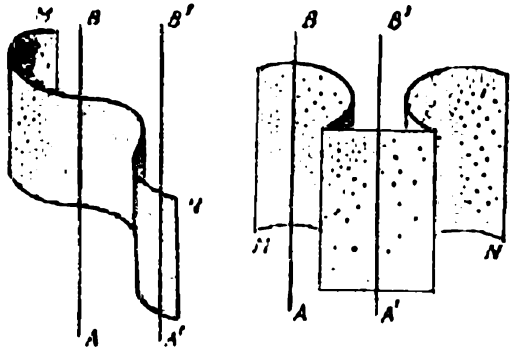


Рис. 5.128

На рис. 5.128 прямые AB и $A'B'$ — образующие, линия MN — направляющая.

Цилиндром называется фигура, ограниченная цилиндрической поверхностью с замкнутой направляющей и двумя параллельными плоскостями, не содержащими образующих цилиндрической поверхности и пересекающими цилиндрическую поверхность.

*) Формула Эйлера справедлива не только для правильных многогранников, но и вообще для всех выпуклых многогранников (например, призмы, пирамиды и т. д.).

Части параллельных плоскостей, ограничивающих цилиндр, называются *основаниями* цилиндра. Отрезок прямой, перпендикулярной основаниям цилиндра с концами, принадлежащими плоскостям оснований, называется *высотой* цилиндра. Отрезки образующих цилиндрической поверхности, заключенные между основаниями цилиндра, называются его *образующими*.

Для любого цилиндра объем и площадь боковой поверхности вычисляются по формулам

$$V = F \cdot h = Ql,$$

$$S_{бок} = P \cdot h = pl,$$

где F — площадь основания, Q — площадь перпендикулярного сечения, h — высота, l — образующая цилиндра, P — периметр основания, p — периметр перпендикулярного сечения (см. рис. 5.129).

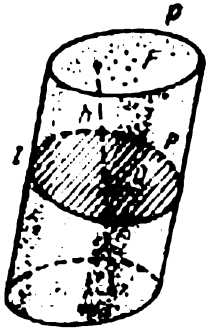


Рис. 5.129

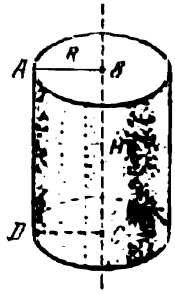


Рис. 5.130

Цилиндр называется *прямым*, если его образующие перпендикулярны основанию; в противном случае — *наклонным*.

Цилиндр называется *прямым круговым*, если образующие перпендикулярны основанию, а основанием цилиндра является круг.

Далее будет рассматриваться только прямой круговой цилиндр (который будем называть просто цилиндром).

Прямой круговой цилиндр может быть также определен как фигура, полученная при вращении прямоугольника вокруг оси, проходящей через одну из его сторон.

При вращении вокруг той же оси ломаной, составленной из сторон прямоугольника, не лежащих на оси вращения, получается фигура, которая называется *поверхностью* цилиндра (рис. 5.130).

Круги, полученные в результате вращения сторон, смежных со стороной, принадлежащей оси вращения, называются *основаниями* цилиндра (на рис. 5.130 основания цилиндра получаются в результате вращения сторон прямоугольника AB и CD вокруг оси BC). Радиус этих двух равных кругов называется *радиусом* основания цилиндра. На рис. 5.130 отрезок AB — радиус основания цилиндра.

Фигура, полученная в результате вращения стороны прямоугольника, противоположной стороне, принадлежащей оси вращения, называется *боковой поверхностью* цилиндра. На рис. 5.130 боковая поверхность цилиндра получается в результате вращения стороны AD вокруг оси BC. Отрезок AD — образующая цилиндра.

Перпендикуляр к плоскостям оснований цилиндра, концы которого совпадают с центрами оснований цилиндра, называется *высотой* цилиндра. На рис. 5.130 отрезок BC — высота цилиндра.

Развертка цилиндра представляет собой прямоугольник, одна сторона которого равна длине окружности основания цилиндра, а другая — высоте цилиндра, и два круга с радиусами, равными радиусу цилиндра.

За *объем цилиндра* принимают предел последовательности объемов правильных призм, вписанных в цилиндр, при бесконечном увеличении числа сторон правильных многоугольников, лежащих в основании призмы *).

Объем цилиндра вычисляется по формуле

$$V = \pi R^2 H,$$

где R — радиус основания цилиндра, H — высота цилиндра.

За *площадь боковой поверхности* цилиндра принимают предел последовательности площадей боковых поверхностей правильных призм, вписанных в цилиндр, при бесконечном увеличении числа сторон правильных многоугольников, лежащих в основании призм.

Площадь боковой и полной поверхности цилиндра вычисляется по формулам

$$S_{\text{бок}} = 2\pi RH,$$
$$S_{\text{цил}} = 2\pi RH + 2\pi R^2,$$

где R — радиус основания цилиндра, H — высота цилиндра.

§ 21. Конус. Усеченный конус

Конической поверхностью называется поверхность, образуемая прямой (образующей), перемещающейся вдоль данной линии (направляющей) и проходящей через данную точку (вершину), не принадлежащую направляющей. На рис. 5.131 AB и $A'B'$ — образующие, MN — направляющая, S — вершина.

Рис. 5.131

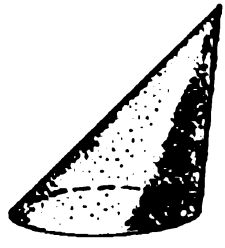


Рис. 5.132

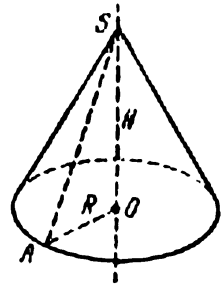


Рис. 5.133

Конусом называется фигура, ограниченная конической поверхностью с замкнутой направляющей и пересекающей ее плоскостью, не проходящей через вершину S. Часть этой плоскости, лежащей внутри конической поверхности, называется *основанием* конуса. Перпендикуляр, опущенный из вершины на плоскость основания, называется *высотой* конуса.

*) О выводе формул вычисления объема цилиндра и площади боковой поверхности цилиндра см. п. 6.4 гл. 8.

Конус называется *круговым*, если его основание — круг (рис. 5.132). Если основание высоты кругового конуса совпадает с центром круга, лежащего в основании конуса, то конус называется *прямым круговым*.

Далее будет рассматриваться только прямой круговой конус (который будем называть просто конусом).

Прямой круговой конус может быть также определен как фигура, полученная при вращении прямоугольного треугольника вокруг оси, содержащей его катет.

Фигура, полученная при вращении вокруг той же оси ломаной, составленной из гипотенузы и катета, не принадлежащего оси вращения, называется *поверхностью* конуса. Фигура, полученная от вращения гипотенузы, называется *боковой поверхностью* конуса, а фигура (круг), полученная от вращения катета,— *основанием* конуса (рис. 5.133). Радиус этого круга называется *радиусом основания* конуса (на рис. 5.133 — отрезок OA).

Катет треугольника, принадлежащий оси, называется *высотой* конуса (на рис. 5.133 отрезок SO — высота конуса). Гипотенуза прямоугольного треугольника называется *образующей* конуса (на рис. 5.133 отрезок AS — образующая конуса).

Развертка боковой поверхности конуса является круговым сектором, а *полная развертка* поверхности конуса представляет собой круговой сектор и круг.

За *объем конуса* принимают предел последовательности правильных пирамид, вписанных в конус, при бесконечном увеличении числа сторон правильного многоугольника — основания пирамиды *).

Объем конуса вычисляется по формуле

$$V_{\text{кон}} = \frac{1}{3}\pi R^2 H,$$

где R — радиус основания конуса, H — высота конуса.

За *площадь боковой поверхности* конуса принимается предел последовательности площадей боковых поверхностей правильных пирамид, вписанных в конус, при бесконечном увеличении числа сторон правильного многоугольника, лежащего в основаниях пирамид.

Площадь боковой поверхности конуса вычисляется по формуле

$$S_{\text{бок}} = \pi RL,$$

где R — радиус основания конуса, L — образующая конуса ($L = |AS|$ на рис. 5.133).

Площадь полной поверхности конуса вычисляется по формуле

$$S_{\text{кон}} = \pi RL + \pi R^2.$$

Усеченный конус. Часть конуса, ограниченная его основанием и сечением, параллельным плоскости основания, называется *усеченным конусом*.

Основания усеченного конуса — гомотетичные круги с центром гомотетии в вершине конуса (на рис. 5.134 — в точке S).

Усеченный конус можно получить в результате вращения равнобедренной трапеции вокруг ее оси симметрии (на рис. 5.134 — трапеции AA_1B_1B).

*) О выводе формул для вычисления объема и площади боковой поверхности конуса см. п. 6.4 гл. 8.

Боковая сторона трапеции называется *образующей* усеченного конуса; круги, полученные при вращении оснований трапеции, — основаниями усеченного конуса.

Развертка усеченного конуса представляет собой объединение части кругового кольца и двух кругов.

Объем усеченного конуса равен разности объемов полного конуса и конуса, отсекаемого плоскостью, параллельной основанию конуса. Объем усеченного конуса вычисляется по формуле

$$V=\frac{1}{3}\pi H\left(R_1^2+R_1R_2+R_2^2\right),$$

где H — высота усеченного конуса (на рис. 5.134 $H=|OO_1|$), R_1 и R_2 — радиусы верхнего и нижнего оснований усеченного конуса (на рис. 5.134 $R_1=|OA|$, $R_2=|O_1A_1|$).

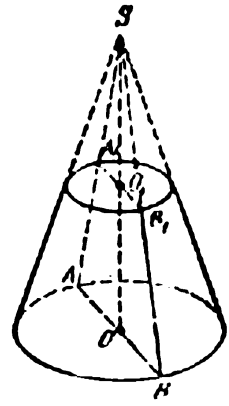


Рис. 5.134

Площадь боковой поверхности усеченного конуса равна разности площадей боковых поверхностей полного конуса и конуса, отсекаемого плоскостью, параллельной основанию конуса.

Площадь боковой поверхности усеченного конуса вычисляется по формуле

$$S_{\text{бок}}=\pi\left(R_1+R_2\right)L,$$

где L — образующая усеченного конуса (на рис. 5.134 $L=|AA_1|$).

§ 22. Сфера. Шар

Множество всех точек пространства, находящихся на данном положительном расстоянии R от данной точки пространства O, называется *сферой*. Данная точка O называется *центром* сферы (рис. 5.135).

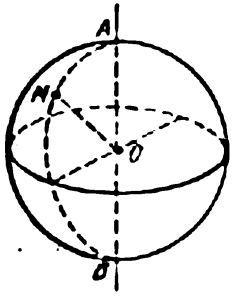


Рис. 5.135

Отрезок OM (M — произвольная точка сферы) называется *радиусом* сферы ($|OM|=R$). Отрезок, соединяющий любые две точки сферы, называется ее *хордой*. Хорда, проходящая через центр сферы, называется *диаметром* сферы. Диаметр сферы равен ее удвоенному радиусу.

Множество всех точек пространства, находящихся от данной точки O на расстоянии, не большем данного расстояния R, называется *шаром*.

Шар можно получить в результате вращения полукруга вокруг оси AB, содержащей диаметр полукруга (см. рис. 5.135). Фигура, полученная при вращении полуокружности, есть сфера — поверхность шара. Центр, радиус и хорды этой сферы называются соответственно *центром*, *радиусом* и *хордами* шара. Все точки шара, не принадлежащие его поверхности, называются *внутренними точками* шара.

Объем шара радиуса R вычисляется по формуле *)

$$V=\frac{4}{3}\pi R^3.$$

Площадь сферы радиуса R вычисляется по формуле

$$S=4\pi R^2.$$

Сечением сферы плоскостью является:

1) окружность, если расстояние от центра сферы до плоскости сечения меньше радиуса сферы;

2) точка, если расстояние от центра сферы до плоскости сечения равно радиусу.

Сечение шара плоскостью, проходящей через центр, называется *большим кругом*. Радиус большого круга равен радиусу шара. На рис. 5.136 L_1 и L_2 — окружности больших кругов.

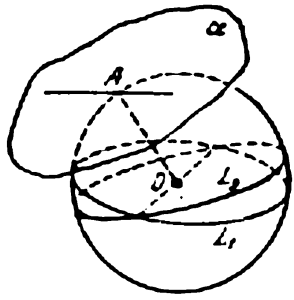


Рис. 5.136

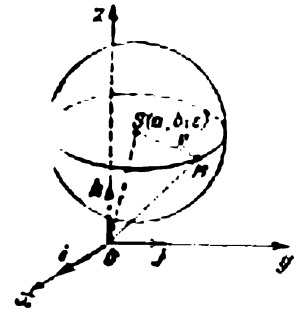


Рис. 5.137

Всякая пара больших кругов пересекается по диаметру шара, являющемуся диаметром для каждого из пересекающихся кругов. Через две точки сферы, лежащие на концах одного и того же диаметра, можно привести бесчисленное множество больших кругов. Через две точки, не лежащие на концах одного диаметра шара, можно провести один и только один большой круг.

Касательной плоскостью к сфере (шару) называется плоскость, имеющая со сферой единственную общую точку. Эту точку (точка A на рис. 5.136) называют *точкой касания* сферы и плоскости. Для того чтобы плоскость была бы касательной к сфере, необходимо и достаточно, чтобы эта плоскость была перпендикулярна радиусу сферы и проходила через его конец. На рис. 5.136 касательная плоскость α перпендикулярна радиусу OA.

Прямая, принадлежащая касательной плоскости к сфере и проходящая через точку касания (точка A на рис. 5.136), называется прямой, *касательной* к сфере (шару). Через каждую точку, принадлежащую сфере, можно провести сколь угодно много касательных.

В прямоугольной пространственной системе координат $Oxyz$ (рис. 5.137) уравнение сферы радиуса R с центром в точке $S(a; b; c)$ имеет вид

$$(x-a)^2+(y-b)^2+(z-c)^2=R^2.$$

*) Вывод формул вычисления объема шара и площади сферы см. в п. 5.4 гл. 9.

§ 23. Части шара

23.1. Шаровой сегмент. Фигура, полученная при вращении кругового сегмента вокруг радиуса, перпендикулярного его хорде, называется *шаровым сегментом* (на рис. 5.138 AB — хорда; CO — радиус).

Фигура, полученная в результате вращения дуги кругового сегмента, называется *сегментной поверхностью*, а фигура, полученная в результате вращения хорды, — *основанием* шарового сегмента. Отрезок радиуса, принадлежащий одновременно оси вращения и круговому сегменту (на рис. 5.138 — отрезок KC), называется *высотой* шарового сегмента (а также высотой сегментной поверхности).

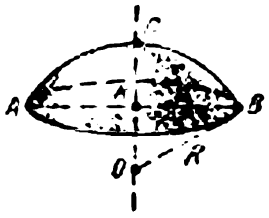


Рис. 5.138

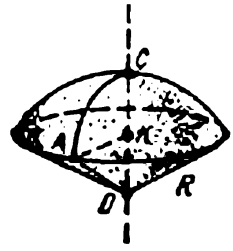


Рис. 5.139

Объем шарового сегмента шара радиуса R и высотой H вычисляется по формуле

$$V=\frac{1}{3}\pi H^2(3R-H).$$

Площадь сегментной поверхности вычисляется по формуле

$$S=2\pi RH.$$

23.2. Шаровой сектор. *Шаровым сектором* называют фигуру, полученную при вращении кругового сектора вокруг оси, проходящей через один из его радиусов (рис. 5.139). Дуга кругового сектора образует при этом вращении сегментную поверхность.

Объем шарового сектора вычисляется по формуе

$$V=\frac{2}{3}\pi R^2H,$$

где R — радиус шара, H — высота шарового сегмента (на рис. 5.139 $H=|KC|$).

Площадь полной поверхности шарового сектора равна сумме площади поверхности шарового сегмента и площади боковой поверхности конуса, полученного при вращении прямоугольного треугольника OKA вокруг оси CO, содержащей его катет (см. рис. 5.139).

Площадь полной поверхности шарового сектора вычисляется по формуле

$$S_{\text{ш. сект}}=S_{\text{сегм}}+S_{\text{бок. кон}}=$$
$$=2\pi RH+\pi R\sqrt{2RH-H^2}=\pi R\left(2H+\sqrt{2RH-H^2}\right).$$

23.3. Шаровой слой. Фигура, полученная при вращении части круга, заключенной между двумя параллельными хордами, вокруг оси, проходящей через диаметр, перпендикулярный хордам, назы-

вается *шаровым слоем* (рис. 5.140). Круги, полученные при вращении хорд, называются *основаниями* шарового слоя. Радиусы этих кругов (на рис. 5.140 — отрезки AD и BC) называются *радиусами шарового слоя*. Отрезок диаметра, принадлежащий одновременно оси вращения и шаровому слою (на рис. 5.140 — отрезок AB), называется *высотой* шарового слоя.

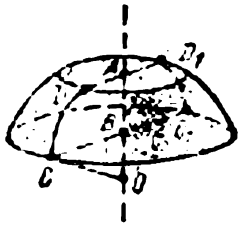


Рис. 5.140

Объем шарового слоя вычисляется по формуле

$$V=\frac{1}{6}\pi H^3+\frac{1}{2}\pi(r_1^2+r_2^2)H,$$

где r_1 и r_2 — радиусы шарового слоя (на рис. 5.140 $|AD|=r_1$, $|BC|=r_2$), H — высота шарового слоя (на рис. 5.140 $H=|AB|$).

23.4. Шаровой пояс. Фигура, полученная в результате вращения дуги окружности, заключенной между двумя параллельными хордами, вокруг оси, проходящей через диаметр, перпендикулярный хордам, называется *шаровым поясом* (или шаровой зоной). На рис. 5.140 шаровой пояс получен в результате вращения дуги DC вокруг оси AB.

Высота шарового слоя одновременно является и высотой соответствующего шарового пояса.

Площадь шарового пояса вычисляется по формуле

$$S_{\text{ш. п}}=2\pi RH,$$

где H — высота шарового пояса, R — радиус дуги, от вращения которой получен шаровой пояс (на рис. 5.140 $R=|CO|$).

23.5. Телесный угол. Пусть L — некоторая замкнутая плоская линия, не имеющая самопересечений, S — некоторая точка, не принадлежащая плоскости линии L (рис. 5.141). Объединение всех лучей, имеющих общее начало S и пересекающих часть плоскости, ограниченную линией L, называется *телесным углом*. Точка S называется *вершиной* телесного угла. Телесный угол измеряется площадью поверхности, вырезаемой телесным углом на сфере радиуса R с центром в вершине телесного угла.

Рис. 5.141

Единица измерения телесного угла называется *стерадианом*.

Один стерадиан — это такой телесный угол, который вырежет на поверхности сферы радиуса R фигуру с площадью, равной R^2.

Многогранный угол является частным случаем телесного угла и также измеряется в стерадианах, например трехгранный угол, образованный в результате пересечения трех координатных плоскостей прямоугольной системы координат $Oxyz$, равен $\pi/2$ стерадиан. Действительно, этот телесный угол высекает $1/8$ часть поверхности шара радиуса R с центром в начале координат и, следовательно, равен

$$\frac{\frac{1}{8}\cdot 4\pi R^2}{R^2}=\frac{\pi}{2}.$$

§ 24. Преобразования плоскости и пространства

24.1. Геометрические фигуры. Отображение фигуры в фигуру и отображение фигуры на фигуру. Геометрической *фигурой* называется любое множество точек пространства (плоскости). Геометрическую фигуру называют плоской, если все ее точки принадлежат одной плоскости.

Пересечение двух (или нескольких) геометрических фигур — это фигура, состоящая из всех тех и только тех точек, которые принадлежат каждой из данных фигур. В теории множеств любые два множества имеют пересечение (которое может быть и пустым множеством). В геометрии обычно вместо слов «пересечение двух фигур пусто» говорят, что фигуры не пересекаются.

Объединение двух (или нескольких) геометрических фигур есть фигура, состоящая из всех тех и только тех точек, которые принадлежат хотя бы одной из данных фигур.

Пусть Φ и Φ_1 — две геометрические фигуры. Если каждой точке M фигуры Φ некоторым способом поставлена в соответствие единственная точка M_1 фигуры Φ_1, то такое соответствие называется *отображением фигуры* Φ *в фигуру* Φ_1. При этом точка M_1 называется *образом* точки M, а точка M — *прообразом* точки M_1. Если при выбранном отображении фигуры Φ в фигуру Φ_1 некоторые точки фигуры Φ_1 имеют несколько прообразов, то множество всех прообразов точки $M_1 \in \Phi_1$ называется *полным прообразом* точки M_1.

Множество Φ' всех образов точек фигуры Φ называется *образом* фигуры Φ.

Если при выбранном отображении f образ фигуры Φ совпадает с фигурой Φ_1, т. е. $\Phi' = \Phi_1$, то говорят, что фигура Φ *отображается на фигуру* Φ_1, и пишут

$$\Phi \xrightarrow{f} \Phi_1 \quad \text{или} \quad f\colon \Phi \to \Phi_1.$$

Отображение фигуры Φ на фигуру Φ_1, при котором разным точкам фигуры Φ соответствуют разные точки фигуры Φ_1, называется *взаимно однозначным* отображением фигуры Φ на фигуру Φ_1.

Пусть f — взаимно однозначное отображение фигуры Φ на фигуру Φ_1. Отображение, ставящее в соответствие каждой точке M_1 фигуры Φ_1 ее прообраз M при отображении f называют *обратным* отображением, обозначают символом f^{-1} и пишут

$$\Phi_1 \xrightarrow{f^{-1}} \Phi \quad \text{или} \quad f^{-1}\colon \Phi_1 \to \Phi.$$

24.2. Преобразование плоскости и пространства. Взаимно однозначное отображение пространства (плоскости) на себя называется *преобразованием* пространства (плоскости).

Простейшие преобразования пространства (плоскости).

1) Преобразования пространства (плоскости), сохраняющие расстояния между любыми двумя точками пространства (плоскости). Такие преобразования называются *изометрией* *).

2) Преобразования пространства (плоскости), при котором расстояния между любыми двумя точками изменяются в одном и том же отношении $k > 0$. Такие преобразования пространства (плоскости) называются преобразованиями *подобия*.

*) Иногда вместо термина «изометрия» употребляется термин «перемещение».

Данное разбиение преобразований на два вида весьма условно — изометрия является частным случаем преобразования подобия; преобразование подобия есть изометрия, если $k=1$.

Пусть f — произвольное преобразование пространства (плоскости). Если преобразование f отображает точку M пространства (плоскости) на точку M_1 пространства (плоскости), то пишут

$$f(M)=M_1.$$

Точка M_1 называется *образом* точки M при преобразовании пространства (плоскости) f, а точка M — *прообразом* точки M_1.

Преобразование пространства (плоскости) f можно представлять как множество всех упорядоченных пар точек $(M; M_1)$ пространства (плоскости), где $M_1=f(M)$. Множество всех упорядоченных пар точек пространства (плоскости) $(M_1; M)$ определяет обратное преобразование. Такое преобразование называется *обратным* преобразованием по отношению к преобразованию f и обозначается символом f^{-1}.

Композиция преобразований. Пусть f_1 и f_2 — два последовательных преобразования пространства (плоскости):

$$M_1=f_1(M) \quad \text{и} \quad M_2=f_2(M_1).$$

Преобразование пространства (плоскости), отображающее точку M пространства (плоскости) в точку M_2 пространства (плоскости), называется *композицией преобразований* f_1 и f_2 и обозначается

$$f_2 \circ f_1.$$

Тождественное преобразование. Преобразование пространства (плоскости), отображающее каждую точку пространства (плоскости) на себя, называется *тождественным* преобразованием пространства (плоскости). Тождественное преобразование обозначается буквой E. Согласно определению тождественного преобразования пространства (плоскости)

$$E(M)=M.$$

Преобразование E можно рассматривать как множество $(M; M)$ всех пар совпадающих точек пространства (плоскости).

24.3. Изометрия пространства и плоскости. Равенство фигур. Из всех преобразований пространства (плоскости) можно выделить преобразования, при которых расстояние между двумя любыми точками пространства (плоскости) равно расстоянию между их образами, т. е. если $f(A)=A_1$ и $f(B)=B_1$, где A и B — любые две точки пространства (плоскости), то $|AB|=|A_1B_1|$. О таких преобразованиях пространства (плоскости) говорят, что они сохраняют расстояние.

Преобразование пространства (плоскости), сохраняющее расстояние, называется *изометрией*. Из определения изометрии следует, что:

1) тождественное преобразование есть изометрия;

2) преобразование, обратное изометрии, есть изометрия;

3) композиция двух изометрий есть изометрия.

Любая изометрия плоскости может быть представлена как последовательное выполнение трех преобразований: параллельного переноса, поворота вокруг точки и осевой симметрии.

Две геометрические фигуры Φ и Φ_1 называют *равными*, если существует изометрия, отображающая фигуру Φ на фигуру Φ_1 *).

*) В курсе геометрии средней школы вместо термина «равенство» употребляется термин «конгруэнтность».

24.4. Поворот плоскости вокруг точки. *Поворотом* плоскости вокруг центра O называется такая изометрия плоскости, при которой:

1) точка O отображается сама на себя;

2) угол между любым лучом OX и соответствующим ему лучом OX_1 (образом луча OX) имеет одну и ту же величину α, называемую *углом поворота*.

Если луч OX совмещается с лучом OX_1 поворотом против часовой стрелки, то направление поворота считается *положительным*; если же луч OX совмещается с лучом OX_1 поворотом по часовой

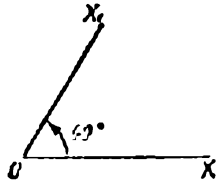


Рис. 5.142

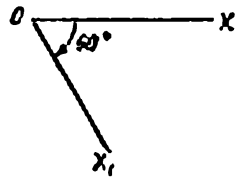


Рис. 5.143

стрелке, то направление поворота считается *отрицательным*. Так, например, на рис. 5.142 лучи OX и OX_1 задают поворот на 60°, а на рис. 5.143 лучи OX и OX_1 задают поворот на $-60°$.

Поворот вокруг центра O с углом поворота α обозначается R_O^α, где $\alpha \in (-180°; 180°]$.

Поворот на угол $\beta = \alpha + 360° \cdot n$, где $n \in Z$ и $\alpha \in (-180°; 180°]$, отождествляется с поворотом R_O^α.

Так как поворот плоскости вокруг точки является частным случаем изометрии, то для него справедливы все общие свойства изометрии:

1) отображение, обратное повороту вокруг центра O, также является поворотом вокруг центра O. Так, например, для поворота против часовой стрелки на угол $\alpha = 30°$ обратным является поворот тоже на 30°, но по часовой стрелке;

2) композиция двух поворотов вокруг центра O на углы α_1 и α_2 есть поворот на угол $\alpha_1 + \alpha_2$: $R_O^{\alpha_1} \circ R_O^{\alpha_2} = R_O^{\alpha_1 + \alpha_2}$, причем композиция двух поворотов обладает свойством коммутативности:

$$R_O^{\alpha_1} \circ R_O^{\alpha_2} = R_O^{\alpha_2} \circ R_O^{\alpha_1}.$$

24.5. Центральная симметрия и центрально симметричные фигуры. Частный случай поворота плоскости вокруг центра O, а именно поворот на 180°, называется *центральной симметрией* с центром O.

Центральная симметрия с центром O обозначается символом Z_O. По определению центральной симметрии

$$Z_O = R_O^{180°}.$$

Геометрическая фигура, которая при центральной симметрии с центром O отображается сама на себя, называется *центрально симметричной* (говорят также, что такая фигура имеет *центр симметрии* O). Центрально симметричная фигура с центром симметрии O изображена на рис. 5.144.

24.6. Осевая симметрия плоскости. Говорят, что точки M_1 и M_2 плоскости α *симметричны* относительно прямой l, принадлежащей

плоскости α и не содержащей точек M_1, M_2, если отрезок M_1M_2 перпендикулярен прямой l и делится этой прямой пополам ($|M_1C| = |OM_2|$, O — середина отрезка M_1M_2 (рис. 5.145)).

Любая точка прямой l считается симметричной самой себе.

Изометрия плоскости, при которой каждая точка плоскости отображается на симметричную ей (относительно данной прямой) точку, называется *осевой симметрией* плоскости.

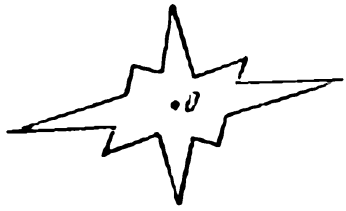


Рис. 5.144

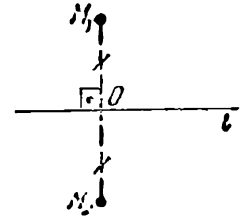


Рис. 5.145

Данную прямую l называют *осью симметрии*, а осевую симметрию относительно данной прямой l обозначают символом S_l.

Так как осевая симметрия есть частный случай изометрии, то для осевой симметрии справедливы все общие свойства изометрии. В частности, отображение, обратное осевой симметрии, есть осевая симметрия.

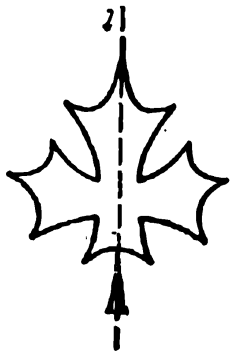


Рис. 5.146

Для осевой симметрии справедливо и более сильное утверждение: отображение, обратное осевой симметрии, есть та же самая осевая симметрия.

Прямая l называется *осью симметрии* фигуры Φ, если фигура Φ при осевой симметрии с осью l отображается сама на себя. В этом случае о фигуре Φ говорят, что она *симметрична* относительно оси l.

На рис. 5.146 изображена фигура, симметричная относительно оси l.

24.7. Осевая симметрия пространства. Говорят, что точки пространства M_1 и M_2 *симметричны* относительно прямой l, не содержащей точек M_1, M_2, если отрезок M_1M_2 перпендикулярен l и делится этой прямой пополам ($|M_1O| = |OM_2|$, где O — середина отрезка M_1M_2, принадлежащая прямой l).

Любая точка прямой l считается симметричной самой себе.

Изометрия пространства, при которой каждая точка пространства отображается на симметричную ей (относительно данной прямой l) точку, называется *осевой симметрией* пространства.

Данную прямую l называют *осью симметрии*. Осевая симметрия пространства (так же как и осевая симметрия плоскости) обозначается символом S_l.

Осевая симметрия пространства является частным случаем изометрии пространства, поэтому для осевой симметрии пространства справедливы все общие свойства изометрии пространства.

Если фигура Φ при осевой симметрии с осью l отображается сама на себя, то говорят, что фигура Φ *симметрична* относительно оси l, а ось l называют *осью симметрии* фигуры Φ.

На рис. 5.147 изображены две пространственные геометрические фигуры с осью симметрии l.

Частным случаем фигуры Ф, симметричной относительно оси l, являются фигуры вращения. Пусть в пространстве дана прямая l и выбрана некоторая произвольная точка M, не принадлежащая пря-

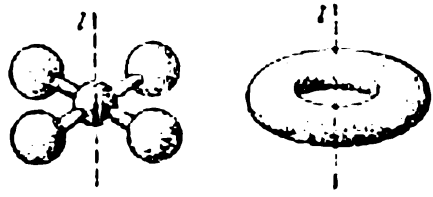


Рис. 5.147

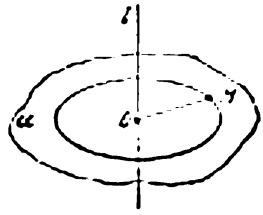


Рис. 5.148

мой l. Построим плоскость α, перпендикулярную прямой l и содержащую точку M (рис. 5.148). Плоскость α пересечет прямую l в некоторой точке O. Рассмотрим окружность радиуса OM с центром в точке O, принадлежащую плоскости α. О такой окружности говорят, что она получена при вращении точки M вокруг оси l.

Теперь рассмотрим в пространстве прямую l и некоторую фигуру F, принадлежащую плоскости, проходящей через данную прямую l (рис. 5.149). Каждой точке фигуры F, не принадлежащей прямой l, поставим в соответствие окружность, получающуюся в результате вращения данной точки вокруг прямой l.

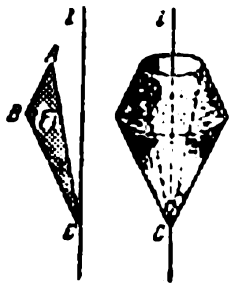


Рис. 5.149

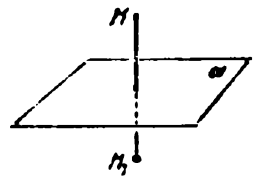


Рис. 5.150

Объединение всех таких окружностей и точек фигуры F, принадлежащих прямой l, называется *фигурой вращения*, полученной в результате вращения фигуры F вокруг прямой l, которая называется *осью вращения*.

Сечением фигуры вращения называется непустое пересечение фигуры вращения и плоскости.

Сечение фигуры вращения называется *осевым* сечением, если оно представляет собой пересечение фигуры вращения плоскостью, проходящей через ось вращения фигуры.

24.8. Симметрия относительно плоскости. Говорят, что точки пространства M и M_1 *симметричны* относительно плоскости α (рис. 5.150), если отрезок MM_1 перпендикулярен этой плоскости и делится ею пополам. Любая точка плоскости α считается симметричной самой себе.

Преобразование пространства, при котором каждая точка отображается на симметричную ей относительно данной плоскости точку,

называется *симметрией* относительно этой плоскости. Симметрию относительно плоскости α обозначают символом S_α.

Плоскость α называется *плоскостью симметрии* фигуры Φ, если при симметрии относительно плоскости α фигура Φ отображается сама на себя. В этом случае о фигуре Φ говорят, что она симметрична относительно плоскости α.

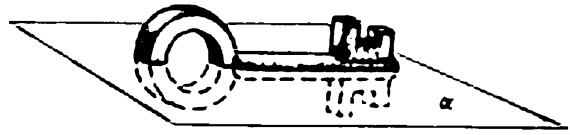


Рис. 5.151

На рис. 5.151 изображена фигура, симметричная относительно плоскости α.

24.9. Гомотетия плоскости. *Гомотетией* плоскости с *центром* O и *коэффициентом* $k \neq 0$ называется отображение плоскости на себя, при котором образом произвольной точки X является точка X_1 такая, что

$$\overrightarrow{OX_1} = k\overrightarrow{OX}.$$

Гомотетию с центром O и коэффициентом k обозначают символом H_O^k:

$$H_0^k(X) = X_1, \quad \text{если} \quad \overrightarrow{OX_1} = k\overrightarrow{OX}.$$

Говоря о множестве гомотетий с каким-либо определенным центром, в обозначении гомотетии букву O опускают, т. е. вместо H_O^k пишут H^k. Гомотетии с отдельно указанными центром и коэффициентом обозначают H.

Основные свойства гомотетии.

1. Центр гомотетии отображается на себя.

2. Если коэффициент $k > 0$, то точки X и $X_1 = H(X)$ лежат на прямой OX по одну сторону от центра гомотетии, т. е. точка X_1 принадлежит лучу OX.

Если $k < 0$, то точки X и $X_1 = H(X)$ лежат на прямой OX по разные стороны от центра гомотетии, т. е. точка X_1 принадлежит лучу, имеющему то же самое начало O, что и луч OX, и противоположно направленному с лучом OX.

3. Если при гомотетии с коэффициентом k точки X и Y отображаются соответственно на точки X_1 и Y_1, то

$$\overrightarrow{X_1Y_1} = k\overrightarrow{XY}.$$

4. При гомотетии с коэффициентом 1 каждая точка переходит сама в себя. Поэтому тождественное отображение плоскости на себя есть гомотетия с любым центром и коэффициентом $k = 1$.

5. При гомотетии с центром O и коэффициентом -1 каждая точка X переходит в точку X_1, для которой $\overrightarrow{OX_1} = -\overrightarrow{OX}$, т. е. в точку, центрально симметричную точке X. Поэтому гомотетия с коэффициентом $k = -1$ есть центральная симметрия:

$$H_0^{-1} = Z_0.$$

6. Обратным отображением к гомотетии с коэффициентом k является гомотетия с тем же центром и коэффициентом $k' = 1/k$.

7. При гомотетии с коэффициентом k все расстояния между точками изменяются в $|k|$ раз:

$$|X_1Y_1| = |k| \cdot |XY|,$$

где $X_1 = H^k(X)$, $Y_1 = H^k(Y)$.

Две плоские фигуры Φ и Φ_1 называются *гомотетичными*, если существует гомотетия, отображающая фигуру Φ на фигуру Φ_1.

Свойства гомотетичных фигур.

1. Гомотетичные фигуры подобны.

2. При гомотетии с положительным коэффициентом каждый луч переходит в сонаправленный с ним луч. При гомотетии с отрицательным коэффициентом каждый луч переходит в противоположно направленный луч. При гомотетии любая прямая, проходящая через центр гомотетии, переходит сама в себя; прямая, не проходящая через центр гомотетии (при $k \neq 1$), переходит в параллельную прямую; угол переходит в равный угол и т. д.

24.10. Гомотетия пространства. Гомотетия пространства определяется так же, как и гомотетия плоскости.

Гомотетией пространства с *центром* O и *коэффициентом* $k \neq 0$ называется преобразование пространства, при котором образом произвольной точки X является такая точка X_1, что $\overrightarrow{OX_1} = k\overrightarrow{OX}$. Гомотетия пространства с центром O и коэффициентом k обозначается тем же символом H_O^k, что и гомотетия плоскости.

Все свойства гомотетии плоскости справедливы и для гомотетии пространства. Кроме того:

1) при гомотетии пространства плоскость отображается либо на саму себя, либо на параллельную плоскость;

2) всякая сфера при гомотетии пространства с центром, совпадающим с центром сферы, отображается на сферу.

Гомотетия как пространства, так и плоскости задается либо центром гомотетии и упорядоченной парой соответствующих точек X и $X_1 = H(X)$, либо двумя упорядоченными парами соответственных точек $(X; X_1)$ и $(Y; Y_1)$, где $X_1 = H(X)$ и $Y_1 = H(Y)$.

Кроме перечисленных выше способов задания, гомотетия может задаваться, например, центром гомотетии O и коэффициентом гомотетии k.

24.11. Преобразование подобия плоскости. Отображение плоскости на себя, при котором расстояния между любыми двумя точками изменяются в одном и том же отношении $k > 0$, называется *преобразованием* подобия или просто *подобием*. Число k называется *коэффициентом* подобия.

Так, если f — преобразование подобия, изменяющее расстояние в k раз, а A и B — две любые точки плоскости и $f(A) = A_1$, $f(B) = B_1$, то $|A_1B_1| = k \cdot |AB|$.

Свойства преобразований подобия.

1. Композиция $f_1 \circ f_2$ двух преобразований подобия f_1 и f_2 с коэффициентами подобия k_1 и k_2 соответственно есть преобразование подобия с коэффициентом $k_1 \cdot k_2$. В частности, композиция гомотетии и изометрии есть преобразование подобия.

2. Каждое преобразование подобия есть композиция гомотетии и изометрии.

3. Каждая гомотетия является преобразованием подобия.

4. Каждая изометрия является преобразованием подобия с коэффициентом подобия, равным единице.

24.12. Подобные фигуры. Если фигуру Φ можно отобразить на фигуру Φ_1 так, что для любых точек A и B фигуры Φ отношение расстояния $|A_1B_1|$ между их образами (точками фигуры Φ_1) к расстоянию $|AB|$ между самими точками равно одному и тому же числу $k(|A_1B_1| = k \cdot |AB|)$, то говорят, что фигура Φ_1 *подобна* фигуре Φ с коэффициентом подобия k, и пишут

$$\Phi_1 \overset{k}{\sim} \Phi.$$

Свойства подобных фигур:

1. Каждая фигура подобна самой себе с коэффициентом подобия $k = 1$ (рефлексивность):

$$\Phi \overset{1}{\sim} \Phi.$$

2. Если фигура Φ_1 подобна фигуре Φ с коэффициентом k, то фигура Φ подобна фигуре Φ_1 с коэффициентом $k' = 1/k$ (симметричность):

$$\Phi_1 \overset{k}{\sim} \Phi \Leftrightarrow \Phi \overset{1/k}{\sim} \Phi_1.$$

Принимая во внимание свойство симметричности, обычно просто говорят, что две фигуры Φ и Φ_1 подобны.

3. Если фигура Φ_1 подобна фигуре Φ с коэффициентом k_1, а фигура Φ_2 подобна фигуре Φ_1 с коэффициентом подобия k_2, то фигура Φ_2 подобна фигуре Φ с коэффициентом $k = k_1 \cdot k_2$ (транзитивность):

$$\Phi_1 \overset{k_1}{\sim} \Phi \quad \text{и} \quad \Phi_2 \overset{k_2}{\sim} \Phi_1 \Rightarrow \Phi_2 \overset{k_1 \cdot k_2}{\sim} \Phi_1.$$

§ 25. Система аксиом и неопределяемых понятий геометрии (по Гильберту)

Основными (неопределяемыми) понятиями геометрии по Гильберту являются три вида объектов:

1) точки, которые будем обозначать A, B, C, ... ;

2) прямые, которые будем обозначать a, b, c, ... ;

3) плоскости, которые будем обозначать α, β, γ, ...

Аксиомы геометрии по Гильберту.

I. Аксиомы принадлежности. Аксиомы этой группы устанавливают отношения принадлежности между основными понятиями геометрии.

I-1. Для любых двух точек A и B существует прямая a, проходящая через эти точки.

I-2. Для любых двух различных точек A и B существует не более одной прямой, проходящей через эти точки.

I-3. На прямой существуют по крайней мере две точки. Существуют по крайней мере три точки, не лежащие на одной прямой.

I-4. Для любых трех точек A, B и C, не лежащих на одной и той же прямой, существует плоскость α, содержащая три данные точки A, B и C. Для любой плоскости всегда существует принадлежащая ей точка.

I-5. Для любых трех точек A, B и C, не лежащих на одной и той же прямой, существует не более одной плоскости, проходящей через эти точки. Эту плоскость часто обозначают ABC, т. е. перечислением точек, через которые проходит плоскость.

I-6. Если две точки A, B прямой a лежат в плоскости α, то всякая точка прямой a лежит в плоскости α.

I-7. Если две плоскости α и β имеют одну общую точку A, то они имеют по крайней мере еще одну общую точку B.

I-8. Существуют по крайней мере четыре точки, не лежащие в одной плоскости.

Аксиомы I-1—I-3 называют плоскостными аксиомами, а остальные аксиомы — пространственными.

II. Аксиомы порядка. Аксиомы этой группы определяют понятие «между», служащее для описания отношения порядка точек на прямой, плоскости и в пространстве.

Точки, лежащие на прямой, находятся в определенных отношениях друг с другом. Для описания этих отношений используется слово «между».

II-1. Если точка B лежит между точками A и C, то A, B и C — три различные точки прямой и точка B лежит также между C и A.

II-2. Для любых двух различных точек A и C, лежащих на прямой AC, существует по крайней мере одна точка B такая, что точка C лежит между A и B.

II-3. Среди любых трех различных точек, лежащих на прямой, существует не более одной точки, лежащей между двумя другими.

Кроме этих аксиом порядка на прямой, вводится еще одна аксиома, определяющая отношение порядка на плоскости. В формулировке этой аксиомы используется понятие отрезка, которое вводится следующим образом.

Пусть A и B — две различные точки прямой a. Множество всех точек прямой a, лежащих между точками A и B, включая A и B, называется *отрезком*, а точки A и B — концами отрезка. Отрезок с концами A, B обозначается AB (или BA). Всякая точка отрезка, лежащая между его концами, называется *внутренней точкой* отрезка.

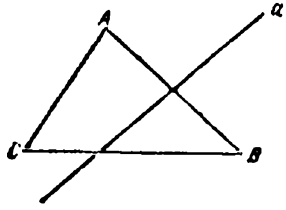


Рис. 5.152

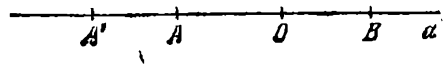


Рис. 5.153

II-4. Пусть A, B и C — три точки, не лежащие на одной прямой, и a — прямая в плоскости ABC, не проходящая ни через одну из точек A, B и C (рис. 5.152). Если при этом прямая a проходит через одну из точек отрезка AB, то она должна пройти через одну из точек отрезка AC или через одну из точек отрезка BC.

Пусть на прямой a заданы четыре различные точки A, B, A' и O (рис. 5.153) так, что точка O лежит между A и B. В этом случае говорят, что точки A и A' прямой a лежат по одну и ту же сторону от точки O, а точки A, B лежат на прямой по разные стороны от точки O. Множество всех точек прямой a, лежащих по одну сторону от точки O, включая точку O, называется *лучом*, исходящим из точки O.

III. Аксиомы конгруэнтности. Отрезки находятся в определенном отношении друг с другом; для описания этого отношения служат понятия конгруэнтности или равенства отрезков *).

III-1. Пусть A и B — две точки прямой a и A' — некоторая точка прямой a' (которая, в частности, может совпадать с прямой a). Тогда на прямой a' сущест-

*) В аксиоматике Гильберта термины «конгруэнтность» и «равенство» являются синонимами.

вует точка B', лежащая по заданную сторону от точки A' и такая, что отрезок AB равен отрезку $A'B'$ (рис. 5.154). Для обозначения равенства отрезков используется знак $=$.

Эта аксиома дает возможность откладывать отрезки.

III-2. Если отрезки $A'B'$ и $A''B''$ равны одному и тому же отрезку AB, то отрезки $A'B'$ и $A''B''$ равны.

Из аксиомы III-2, в частности, вытекает, что любой отрезок равен самому себе, а также что отношение равенства отрезков симметрично и транзитивно.

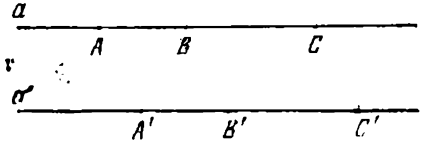


Рис. 5.154

III-3. Пусть AB и BC — два отрезка прямой a, не имеющие ни одной общей внутренней точки, а $A'B'$ и $B'C'$ — два отрезка той же или другой прямой a', также не имеющие общих внутренних точек (см. рис. 5.154). Если при этом

$$AB = A'B' \quad \text{и} \quad BC = B'C',$$

то и $AC = A'C'$.

Эта аксиома дает возможность складывать отрезки.

Аксиомы III-1 и III-3 содержат утверждения, касающиеся лишь равенства отрезков; их называют *линейными аксиомами равенства*.

При формулировке следующих двух аксиом равенства используется понятие угла. Углы находятся в определенном отношении, для обозначения которого используется понятие равенства.

III-4. Пусть даны угол $\angle (h_1, h_2)$ в плоскости α и прямая a' в плоскости α', а также выбрана вполне определенная по отношению к прямой a' часть плоскости α', отделяемая прямой a' от всей плоскости. Пусть h_1' — луч прямой a', исходящий из точки O'. Тогда в плоскости α' существует один и только один луч h_2' такой, что угол $\angle (h_1', h_2')$ равен углу $\angle (h_1, h_2)$ и все внутренние точки угла $\angle (h_1', h_2')$ находятся в плоскости α' по данную сторону от прямой a';

$$\angle (h_1, h_2) = \angle (h_1', h_2').$$

III-5. Если для двух треугольников ABC и $A'B'C'$ имеют место равенства

$$AB = A'B', \quad AC = A'C', \quad \angle BAC = \angle B'A'C',$$

то имеет место также равенство

$$\angle ABC = \angle A'B'C'.$$

Аксиомы III-4 и III-5 называются *плоскостными аксиомами равенства*.

IV. Аксиома о параллельных прямых (аксиома Евклида). Пусть a — произвольная прямая, а A — точка, не принадлежащая прямой a. Тогда в плоскости, определяемой прямой a и точкой A, существует не более одной прямой, проходящей через точку A и не пересекающей прямую a.

Эта прямая называется прямой, *параллельной* a и проходящей через точку A.

Иногда при перечислении аксиом геометрии вместо аксиомы о параллельных выдвигаются предложения, равносильные этой аксиоме. Такими предложениями, например, будут следующие:

1) если две прямые a и b, лежащие в одной плоскости, не пересекают третью прямую, лежащую в той же плоскости, то они не пересекаются также и между собой;

2) если принята аксиома Архимеда (см. ниже), то аксиома о параллельных может быть заменена требованием равенства суммы внутренних углов треугольника двум прямым углам.

V. Аксиомы непрерывности.

V-1 (аксиома измерения, или аксиома Архимеда). Пусть AB и CD — два произвольных отрезка. Тогда на прямой AB существует конечное число точек A_1, A_2, A_3, ..., A_n таких, что все отрезки AA_1, A_1A_2,

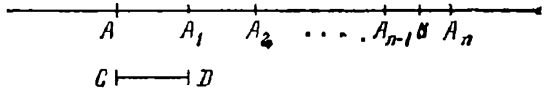


Рис. 5.155

A_2A_3, ..., $A_{n-1}A_n$ равны отрезку CD и точка B лежит между точками A и A_n (рис. 5.155).

V-2 (аксиома линейной полноты). Точки прямой образуют систему, которая (при условии принятия перечисленных аксиом) не допускает никакого расширения, т. е. к множеству точек прямой невозможно добавить еще некоторые объекты так, чтобы:

1) все элементы полученного расширенного множества снова подчинялись перечисленным аксиомам;

2) все аксиомы сохраняли свой первоначальный смысл, т. е. существовавшие до расширения отношение порядка между точками и равенство отрезков сохранились бы и в расширенном множестве.

Глава 6

ТРИГОНОМЕТРИЯ

Тригонометрия возникла как аппарат для вычисления неизвестных параметров треугольника по заданным значениям других его параметров. Так, методами тригонометрии по данным сторонам треугольника можно вычислить его углы, по известной площади и двум углам вычислить стороны и т. д. Необходимость отыскивать неизвестные параметры данного треугольника впервые возникла в астрономии, и в течение долгого времени тригонометрия была одним из ее разделов.

Первые методы нахождения неизвестных параметров данного треугольника были развиты учеными Древней Греции за несколько веков до новой эры. Греческие астрономы не рассматривали синусов, косинусов и тангенсов. Вместо таблиц этих величин они составили и использовали таблицы, позволяющие отыскивать хорду окружности по стягиваемой ею дуге. Дальнейшее развитие тригонометрия получила в средние века в работах индийских и арабских ученых. Современные буквенные обозначения появились в тригонометрии в середине XVIII века. Приблизительно в то же время в тригонометрии стала рассматриваться радианная мера угла, были введены тригонометрические и обратные тригонометрические функции числового аргумента, после чего тригонометрия приобрела свой современный вид.

§ 1. Тригонометрические функции

1.1. Обобщение понятия угла. Пусть $\angle(a, b)$ и $\angle(c, d)$ — два тупых угла, образованных парами лучей a, b и c, d. Совместим лучи b и d так, чтобы они совпали, а лучи a и c лежали по разные стороны от прямой, содержащей лучи b и d (рис. 6.1). В результате сложения этих двух углов получается угол, больший развернутого. Сторонами этого угла будут лучи a и c. Для того чтобы отличать полученный угол от угла (a, c), имеющего те же самые стороны и меньшего развернутого, отметим на рисунке полученный угол дугой. Внутренней областью отмеченного угла будет внешняя область угла $\angle(a, c)$, а внешней — внутренняя область угла $\angle(a, c)$. Аналогичным способом можно получать углы, большие двух развернутых углов (т. е. большие полного угла), и т. д.

Для того чтобы представить себе угол, больший полного, удобно рассматривать угол как геометрическую фигуру, полученную при повороте луча вокруг своего начала.

Возьмем на плоскости луч OA и будем поворачивать его вокруг точки O против часовой стрелки (рис. 6.2). При повороте луча OA на четверть полного оборота вокруг точки O получается угол AOA' (OA' — образ луча OA при указанном повороте), равный 90° ($\pi/2$). Далее, угол AOA' — развернутый, если луч OA' получен поворотом

луча OA на половину полного оборота вокруг точки O. Полный угол получится, когда в результате вращения луч OA' первый раз совпадет с лучом OA. Угол AOA' считается больше полного, если луч OA'

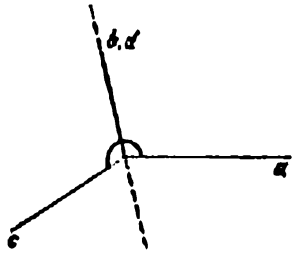


Рис. 6.1

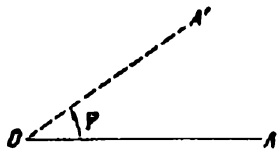


Рис. 6.2

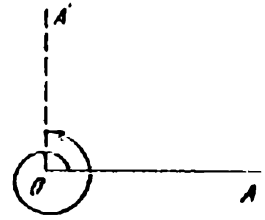


Рис. 6.3

получается поворотом луча OA на угол, больший полного угла. Углы, большие развернутого, часто отмечают, указывая стрелкой направление поворота и число сделанных оборотов. Так, на рис. 6.3 изображен угол в 450°.

Луч OA можно вращать вокруг точки O в двух направлениях: по часовой стрелке и против часовой стрелки. Направление вращения против часовой стрелки называют *положительным*, по часовой стрелке — *отрицательным*. Соответственно углы, полученные вращением луча против часовой стрелки, называют *положительными*, а по часовой стрелке — *отрицательными*.

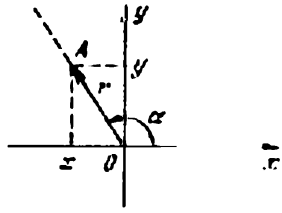


Рис. 6.4

1.2. Тригонометрические функции. Пусть Oxy — прямоугольная декартова система координат; OA — луч, полученный поворотом положительной полуоси Ox вокруг начала координат O на угол α (рис. 6.4) (A — произвольная точка луча). Обозначим координаты точки A через $(x; y)$ и рассмотрим вектор $\overrightarrow{OA} = (x; y)$. Модуль этого вектора равен

$$r = \sqrt{x^2 + y^2}.$$

Углом между осью Ox и вектором $\overrightarrow{OA}$ называют угол между положительным направлением оси Ox и лучом OA.

Отношения

$$\frac{x}{r} \quad \text{и} \quad \frac{y}{r}$$

координат вектора $\overrightarrow{OA}$ к длине вектора $\overrightarrow{OA}$ не зависят от длины вектора $\overrightarrow{OA}$ (т. е. положения точки A на луче OA), но зависят от его направления *).

*) Из множества всех векторов с началом в точке O можно выделить подмножество векторов с началом в точке O и модулем, равным единице. Концы всех таких векторов принадлежат окружности с центром в начале координат и с радиусом, равным единице, — единичной окружности. Тригонометриче-

Отношение ординаты вектора, образующего с осью Ox угол α, к длине этого вектора называется *синусом* угла α и обозначается $\sin\alpha$:

$$\sin\alpha=\frac{y}{r}.$$

Отношение абсциссы вектора, образующего с осью Ox угол α, к длине этого вектора называется *косинусом* угла α и обозначается $\cos\alpha$:

$$\cos\alpha=\frac{x}{r}.$$

Отношение синуса угла α к его косинусу называется *тангенсом* угла α и обозначается $\operatorname{tg}\alpha$:

$$\operatorname{tg}\alpha=\frac{\sin\alpha}{\cos\alpha},$$

или, по определению $\sin\alpha$ и $\cos\alpha$,

$$\operatorname{tg}\alpha=\frac{y}{x}.$$

Отношение косинуса угла α к его синусу называется *котангенсом* угла α и обозначается $\operatorname{ctg}\alpha$:

$$\operatorname{ctg}\alpha=\frac{\cos\alpha}{\sin\alpha},$$

или, по определению $\sin\alpha$ и $\cos\alpha$,

$$\operatorname{ctg}\alpha=\frac{x}{y}.$$

Синус и косинус однозначно определены для любого угла α, тангенс однозначно определен для всех значений α, за исключением $\alpha=90°+180°\cdot n$ $(\alpha=\pi/2+\pi n)$, $n\in Z$, а котангенс — для всех значений α, за исключением $\alpha=180°\cdot n$ $(\alpha=\pi n)$, $n\in Z$.

Синус и косинус одного и того же угла связаны равенством

$$\sin^2\alpha+\cos^2\alpha=1,$$

тангенс и косинус одного и того же угла — равенством

$$1+\operatorname{tg}^2\alpha=\frac{1}{\cos^2\alpha},$$

котангенс и синус — равенством

$$1+\operatorname{ctg}^2\alpha=\frac{1}{\sin^2\alpha},$$

а тангенс и котангенс одного и того же угла — равенством

$$\operatorname{tg}\alpha\cdot\operatorname{ctg}\alpha=1.$$

Функции $\sin\alpha$, $\cos\alpha$, $\operatorname{tg}\alpha$ и $\operatorname{ctg}\alpha$ называют *тригонометрическими функциями* угла α. Кроме основных тригонометрических функций

ские функции часто вводятся с использованием единичной окружности (или множества векторов $\overrightarrow{OA}$ с концами, принадлежащими единичной окружности). Так, если точка A принадлежит единичной окружности, то *синусом* угла α называют ординату вектора $\overrightarrow{OA}$, образующего с осью Ox угол α.

$\sin\alpha$, $\cos\alpha$, $\operatorname{tg}\alpha$ и $\operatorname{ctg}\alpha$, вводятся еще две тригонометрические функции угла α — *секанс* и *косеканс*, обозначаемые $\sec\alpha$ и $\operatorname{cosec}\alpha$ соответственно. Эти функции определяются равенствами

$$\sec\alpha=\frac{1}{\cos\alpha},\quad \operatorname{cosec}\alpha=\frac{1}{\sin\alpha}.$$

Секанс угла α определен при всех значениях α, за исключением $\alpha=90^\circ+180^\circ\cdot n\ \left(\alpha=\frac{\pi}{2}+\pi n\right)$, $n\in Z$, а косеканс — при всех значениях α, за исключением $\alpha=180^\circ\cdot n$ $(\alpha=\pi n)$, $n\in Z$.

1.3. Квадранты единичной окружности. Знаки значений тригонометрических функций. Координатные оси Ox, Oy прямоугольной системы координат делят координатную плоскость на четыре части,

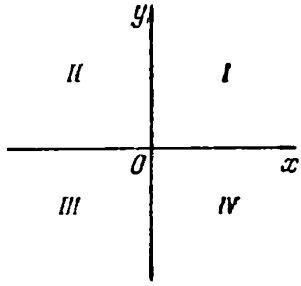


Рис. 6.5

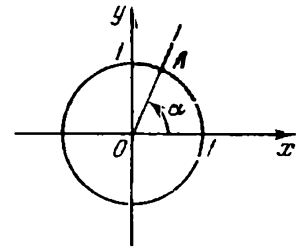


Рис. 6.6

называемые *четвертями* или *квадрантами* (рис. 6.5). Часть плоскости, лежащая выше оси Ox и правее оси Oy, — первая четверть; выше оси Ox и левее оси Oy — вторая; ниже оси Ox и левее оси Oy — третья; ниже оси Ox и правее оси Oy — четвертая. Единичная окружность (окружность единичного радиуса с центром в начале координат) осями координат делится на четыре части, которые тоже называют четвертями (или квадрантами) и нумеруют так же, как и четверти плоскости.

О точках единичной окружности, принадлежащих осям координат, говорят, что они находятся на положительной (или на отрицательной) полуоси абсцисс или ординат.

Пусть луч OA, полученный поворотом положительной полуоси Ox на угол α, пересекает единичную окружность в точке A (рис. 6.6). Угол α считают принадлежащим первой, второй, третьей или четвертой четверти, если точка A принадлежит соответственно первой, второй, третьей или четвертой четверти единичной окружности.

Абсциссы точек, принадлежащих I и IV четвертям, положительны, а точек, принадлежащих II и III четвертям, отрицательны. Ординаты точек, принадлежащих I и II четвертям, положительны, а III и IV четвертям — отрицательны. Знаки значений тригонометрических функций устанавливаются на основании определения тригонометрических функций и зависят от того, какой четверти принадлежит угол α. В табл. 1 приведены знаки значений основных тригонометрических функций.

1.4. Тригонометрические функции числового аргумента. Углы могут измеряться в градусах и радианах. Использование радианной

Таблица 1

Четверти	Функции			
	$\sin\alpha$	$\cos\alpha$	$\operatorname{tg}\alpha$	$\operatorname{ctg}\alpha$
I	+	+	+	+
II	+	−	−	−
III	−	−	+	+
IV	−	+	−	−

меры измерения углов позволяет ввести тригонометрические функции числового аргумента.

Пусть Oxy — прямоугольная декартова система координат. Между множеством всех действительных чисел и множеством всех углов, получаемых в результате поворота луча OA вокруг точки O, можно установить взаимно однозначное соответствие. Будем считать, что положительному числу a соответствует луч OA, полученный в результате поворота положительной полуоси Ox против движения часовой стрелки на a радиан, а отрицательному числу a — луч OA, полученный в результате поворота на a радиан по движению часовой стрелки.

Например, числу 1 соответствует поворот луча OA на один радиан против движения часовой стрелки; числу $-3/2$ соответствует поворот луча OA на $3/2$ радиана по движению часовой стрелки и т. д.

Синусом числа x называется число, равное синусу угла в x радиан. *Косинусом* числа x называется число, равное косинусу угла в x радиан. Аналогично определяются и другие тригонометрические функции числового аргумента.

Основные свойства тригонометрических функций.

Свойства функции $\sin x$.

1. Область определения — множество всех действительных чисел.
2. Область изменения (множество значений) — промежуток $[-1; 1]$.
3. Функция $\sin x$ нечетная: $\sin(-x) = -\sin x$.
4. Функция $\sin x$ периодическая. Наименьший положительный период равен 2π:

$$\sin(x+2\pi) = \sin x.$$

5. Нули функции: $\sin x = 0$ при $x = \pi n$, $n \in \mathbf{Z}$.
6. Промежутки знакопостоянства:

$$\sin x > 0 \text{ при } x \in (2\pi n;\ \pi + 2\pi n),\ n \in \mathbf{Z},$$
$$\sin x < 0 \text{ при } x \in (\pi + 2\pi n;\ 2\pi + 2\pi n),\ n \in \mathbf{Z}.$$

7. Функция $\sin x$ непрерывна и имеет производную при любом значении аргумента:

$$(\sin x)' = \cos x.$$

8. Функция $\sin x$ возрастает при $x \in (-\pi/2 + 2\pi n;\ \pi/2 + 2\pi n)$, $n \in \mathbf{Z}$, и убывает при $x \in (\pi/2 + 2\pi n;\ 3\pi/2 + 2\pi n)$, $n \in \mathbf{Z}$.
9. Функция $\sin x$ принимает минимальные значения, равные -1, при $x = -\pi/2 + 2\pi n$, $n \in \mathbf{Z}$, и максимальные значения, равные 1, при $x = \pi/2 + 2\pi n$, $n \in \mathbf{Z}$.

График функции $y=\sin x$ изображен на рис. 6.7. График функции $\sin x$ называют *синусоидой*.

Свойства функции $\cos x$.

1. Область определения — множество всех действительных чисел.
2. Область изменения (множество значений) — промежуток $[-1; 1]$.
3. Функция $\cos x$ четная: $\cos(-x)=\cos x$.

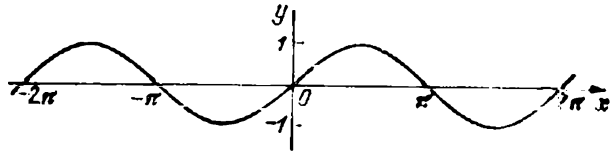


Рис. 6.7

4. Функция $\cos x$ периодическая. Наименьший положительный период функции равен 2π:

$$\cos(x+2\pi)=\cos x.$$

5. Нули функции: $\cos x=0$ при $x=\pi/2+\pi n$, $n\in\mathbf{Z}$.
6. Промежутки знакопостоянства:

$$\cos x>0 \text{ при } x\in(-\pi/2+2\pi n;\ \pi/2+2\pi n),\quad n\in\mathbf{Z},$$
$$\cos x<0 \text{ при } x\in(\pi/2+2\pi n;\ 3\pi/2+2\pi n),\quad n\in\mathbf{Z}.$$

7. Функция $\cos x$ непрерывна и дифференцируема при любом значении аргумента:

$$(\cos x)'=-\sin x.$$

8. Функция $\cos x$ возрастает при $x\in(-\pi+2\pi n;\ 2\pi n)$, $n\in\mathbf{Z}$, и убывает при $x\in(2\pi n;\ \pi+2\pi n)$, $n\in\mathbf{Z}$.

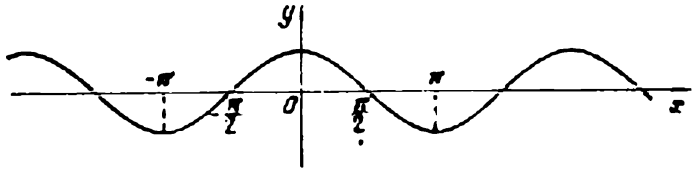


Рис. 6.8

9. Функция $\cos x$ принимает минимальные значения, равные -1, при $x=\pi+2\pi n$, $n\in\mathbf{Z}$, и максимальные значения, равные 1, при $x=2\pi n$, $n\in\mathbf{Z}$.

График функции $y=\cos x$ изображен на рис. 6.8.

Свойства функции $\operatorname{tg} x$.

1. Область определения функции — множество всех действительных чисел, кроме чисел $x=\pi/2+\pi n$, $n\in\mathbf{Z}$.
2. Область изменения (множество значений) — множество всех действительных чисел.
3. Функция $\operatorname{tg} x$ нечетная: $\operatorname{tg}(-x)=-\operatorname{tg} x$.
4. Функция $\operatorname{tg} x$ периодическая. Наименьший положительный период функции равен π:

$$\operatorname{tg}(x+\pi)=\operatorname{tg} x.$$

5. Нули функции: $\operatorname{tg} x=0$ при $x=\pi n$, $n\in\mathbf{Z}$,

6. Промежутки знакопостоянства:

$$\operatorname{tg} x > 0 \text{ при } x \in (\pi n;\ \pi/2+\pi n), \quad n \in \mathbf{Z},$$
$$\operatorname{tg} x < 0 \text{ при } x \in (-\pi/2+\pi n;\ \pi n), \quad n \in \mathbf{Z}.$$

7. Функция $\operatorname{tg} x$ непрерывна и дифференцируема при любом значении аргумента из области определения функции:

$$(\operatorname{tg} x)' = \frac{1}{\cos^2 x}.$$

8. Функция $\operatorname{tg} x$ возрастает в каждом из промежутков

$$(-\pi/2+\pi n;\ \pi/2+\pi n), \quad n \in \mathbf{Z}.$$

График функции $y = \operatorname{tg} x$ изображен на рис. 6.9. График функции $\operatorname{tg} x$ называют *тангенсоидой*.

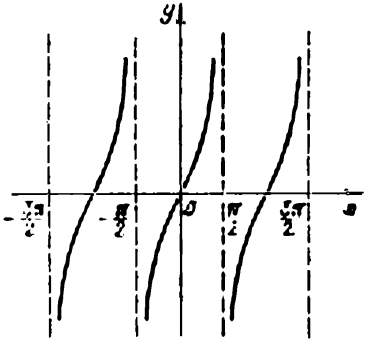


Рис. 6.9

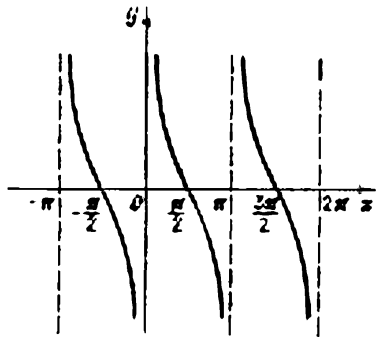


Рис. 6.10

Свойства функции $\operatorname{ctg} x$.

1. Область определения функции — множество всех действительных чисел, кроме чисел $x = \pi n$, $n \in \mathbf{Z}$.

2. Область изменения (множество значений) — множество всех действительных чисел.

3. Функция $\operatorname{ctg} x$ нечетная: $\operatorname{ctg}(-x) = -\operatorname{ctg} x$.

4. Функция $\operatorname{ctg} x$ периодическая. Наименьший положительный период функции равен π:

$$\operatorname{ctg}(x+\pi) = \operatorname{ctg} \pi.$$

5. Нули функции: $\operatorname{ctg} x = 0$ при $x = \pi/2+\pi n$, $n \in \mathbf{Z}$.

6. Промежутки знакопостоянства:

$$\operatorname{ctg} x > 0 \text{ при } x \in (\pi n;\ \pi/2+\pi n), \quad n \in \mathbf{Z},$$
$$\operatorname{ctg} x < 0 \text{ при } x \in (\pi/2+\pi n;\ \pi(n+1)), \quad n \in \mathbf{Z}.$$

7. Функция $\operatorname{ctg} x$ непрерывна и дифференцируема при любом значении аргумента из области определения функции:

$$(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x}.$$

8. Функция $\operatorname{ctg} x$ убывает в каждом из промежутков $(\pi n;\ \pi(n+1))$, $n \in \mathbf{Z}$.

График функции $y = \operatorname{ctg} x$ изображен на рис. 6.10.

Свойства функции $\sec x$.

1. Область определения функции — множество всех действительных чисел, кроме чисел $x=\pi/2+\pi n$, $n\in Z$.

2. Область изменения (множество значений):

$$(-\infty, -1]\cup[1; +\infty).$$

3. Функция $\sec x$ четная: $\sec(-x)=\sec x$.

4. Функция $\sec x$ периодическая. Наименьший положительный период функции равен 2π: $\sec(x+2\pi)=\sec x$.

5. Функция $\sec x$ ни при каком значении аргумента не обращается в нуль.

6. Промежутки знакопостоянства:

$$\begin{aligned}&\sec x>0 \text{ при } x\in(-\pi/2+2\pi n;\ \pi/2+2\pi n), && n\in Z,\\ &\sec x<0 \text{ при } x\in(\pi/2+2\pi n;\ 3\pi/2+2\pi n), && n\in Z.\end{aligned}$$

7. Функция $\sec x$ непрерывна и дифференцируема при любом значении аргумента из области определения функции:

$$(\sec x)'=\frac{\sin x}{\cos^2 x}.$$

8. Функция $\sec x$ возрастает в промежутках

$$[2\pi n;\ \pi/2+2\pi n);\ (\pi/2+2\pi n;\ \pi+2\pi n],\ n\in Z,$$

и убывает в промежутках $[\pi+2\pi n;\ 3\pi/2+2\pi n)$, $(3\pi/2+2\pi n;\ 2\pi(n+1)]$, $n\in Z$.

График функции $y=\sec x$ изображен на рис. 6.11.

Свойства функции $\operatorname{cosec} x$.

1. Область определения функции — множество всех действительных чисел, кроме чисел $x=\pi n$, $n\in Z$.

2. Область изменения (множество значений) функции:

$$(-\infty;\ -1]\cup[1;\ +\infty).$$

3. Функция $\operatorname{cosec} x$ нечетная: $\operatorname{cosec}(-x)=-\operatorname{cosec} x$.

4. Функция $\operatorname{cosec} x$ периодическая. Наименьший положительный период функции равен 2π:

$$\operatorname{cosec}(x+2\pi)=\operatorname{cosec} x.$$

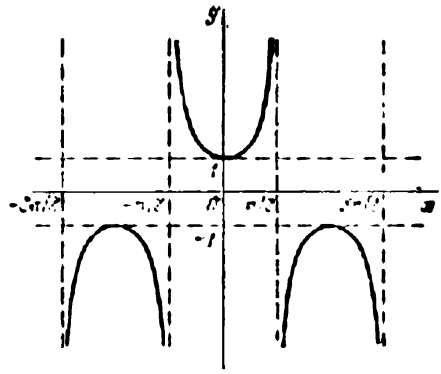


Рис. 6.11

5. Функция $\operatorname{cosec} x$ ни при каком значении аргумента не обращается в нуль.

6. Промежутки знакопостоянства:

$$\begin{aligned}&\operatorname{cosec} x>0 \text{ при } x\in(2\pi n;\ \pi+2\pi n), && n\in Z,\\ &\operatorname{cosec} x<0 \text{ при } x\in(\pi+2\pi n;\ 2\pi(n+1)), && n\in Z.\end{aligned}$$

7. Функция $\operatorname{cosec} x$ непрерывна и дифференцируема при любом значении аргумента из области определения функции:

$$(\operatorname{cosec} x)'=-\frac{\cos x}{\sin^2 x}.$$

8. Функция $\operatorname{cosec} x$ возрастает в промежутках

$$[\pi/2+2\pi n;\ \pi+2\pi n) \quad \text{и} \quad (\pi+2\pi n;\ 3\pi/2+2\pi n], \quad n\in\mathbf{Z},$$

и убывает в промежутках $(2\pi n;\ \pi/2+2\pi n]$ и $[3\pi/2+2\pi n;\ 2\pi+2\pi n)$, $n\in\mathbf{Z}$.

График функции $y=\operatorname{cosec} x$ изображен на рис. 6.12.

1.5. Обратные тригонометрические функции.

Функция $\arcsin x$. Функция $y=\sin x$ определена и непрерывна при любом $x\in\mathbf{R}$. Выделим на числовой оси Ox промежуток $[-\pi/2;\ \pi/2]$. На этом промежутке функция $y=\sin x$ возрастает; на левом конце промежутка (при $x=-\pi/2$) функция $y=\sin x$ достигает своего наименьшего значения, равного -1, а на правом конце (при $x=\pi/2$) — своего наибольшего значения, равного 1. Функция $y=\sin x$, рассматриваемая на промежутке $[-\pi/2;\ \pi/2]$, имеет обратную функцию, которую называют *арксинусом* и записывают

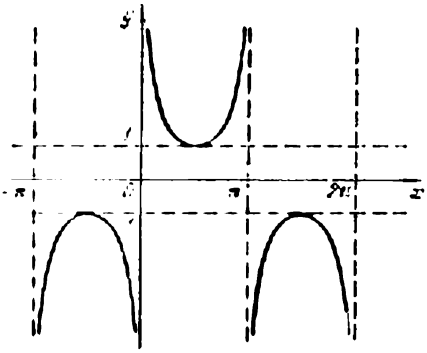


Рис. 6.12

$$x=\arcsin y,$$

где y — независимая переменная, а x — зависимая. Обозначая, как обычно, независимую переменную буквой x, а зависимую — буквой y, будем далее писать $y=\arcsin x$.

Свойства функции $\arcsin x$.

1. Область определения — промежуток $[-1;\ 1]$.

2. Область изменения (множество значений) — промежуток $[-\pi/2;\ \pi/2]$.

3. Функция $\arcsin x$ нечетная: $\arcsin(-x)=-\arcsin x$.

4. Нули функции: $\arcsin x=0$ при $x=0$.

5. Промежутки знакопостоянства:

$$\arcsin x>0 \quad \text{при} \quad x\in(0;\ 1],$$
$$\arcsin x<0 \quad \text{при} \quad x\in[-1;\ 0).$$

6. Функция $\arcsin x$ непрерывна и дифференцируема в каждой точке промежутка $(-1;\ 1)$:

$$(\arcsin x)'=\frac{1}{\sqrt{1-x^2}}.$$

7. Функция $\arcsin x$ возрастает на промежутке $[-1;\ 1]$, принимая свое наименьшее значение, равное $-\pi/2$, на левом конце промежутка и наибольшее значение, равное $\pi/2$, на правом конце промежутка $[-1;\ 1]$.

График функции $x=\arcsin x$ изображен на рис. 6.13.

Функция $\arccos x$. Функция $y=\cos x$ определена и непрерывна при любом $x\in\mathbf{R}$. Выделим на числовой оси Ox промежуток $[0;\ \pi]$. На этом промежутке функция $y=\cos x$ убывает; на левом конце промежутка (при $x=0$) функция $y=\cos x$ достигает своего наибольшего значения, равного 1, а на правом конце (при $x=\pi$) — своего наименьшего значения, равного -1. Функция $y=\cos x$, рассматривая на промежутке $[0;\ \pi]$, имеет обратную функцию, которую

называют *арккосинусом* и записывают

$$x = \arccos y,$$

где y — независимая переменная, а x — зависимая Обозначая, как обычно, независимую переменную буквой x, а зависимую — буквой y, будем далее писать $y = \arccos x$.

Свойства функции $\arccos x$.

1. Область определения — промежуток $[-1; 1]$.
2. Область изменения (множество значений) — промежуток $[0; \pi]$.
3. Функция $\arccos x$ не является ни четной, ни нечетной.
4. Нули функции: $\arccos x = 0$ при $x = 1$.
5. Промежутки знакопостоянства: $\arccos x > 0$ при $x \in [-1; 1]$.
6. Функция $\arccos x$ непрерывна и дифференцируема в каждой точке промежутка $(-1; 1)$;

$$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}.$$

7. Функция $\arccos x$ убывает на промежутке $[-1; 1]$, принимая наибольшее значение, равное π, на левом конце промежутка и наименьшее значение, равное 0, на правом конце промежутка.

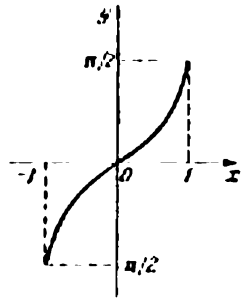


Рис. 6.13

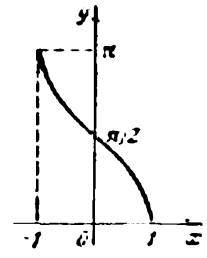


Рис. 6.14

График функции $y = \arccos x$ изображен на рис. 6.14.

Функция $\operatorname{arctg} x$. Функция $y = \operatorname{tg} x$ непрерывна в своей области определения (т. е. при всех $x \in \mathbf{R}$, не равных $\pi/2 + \pi n$, $n \in \mathbf{Z}$). Выделим на числовой оси Ox промежуток $(-\pi/2; \pi/2)$. На этом промежутке функция $y = \operatorname{tg} x$ возрастает и принимает все свои значения. Функция $y = \operatorname{tg} x$, рассматриваемая на промежутке $(-\pi/2; \pi/2)$, имеет обратную функцию, которую называют *арктангенсом* и записывают

$$x = \operatorname{arctg} y,$$

где y — независимая переменная, а x — зависимая. Обозначая, как обычно, независимую переменную буквой x, а зависимую — буквой y, будем далее писать $y = \operatorname{arctg} x$.

Свойства функции $\operatorname{arctg} x$.

1. Область определения — вся числовая прямая.
2. Область изменения (множество значений) — промежуток $(-\pi/2; \pi/2)$.
3. Функция $\operatorname{arctg} x$ нечетная: $\operatorname{arctg}(-x) = -\operatorname{arctg} x$.
4. Нули функции: $\operatorname{arctg} x = 0$ при $x = 0$.

5. Промежутки знакопостоянства:

$$\operatorname{arctg} x > 0 \quad \text{при} \quad x \in (0; +\infty),$$
$$\operatorname{arctg} x < 0 \quad \text{при} \quad x \in (-\infty; 0).$$

6. Функция $\operatorname{arctg} x$ непрерывна и дифференцируема при всех $x \in \mathbf{R}$:

$$(\operatorname{arctg} x)' = \frac{1}{1+x^2}.$$

7. Функция $\operatorname{arctg} x$ возрастающая.

График функции $y = \operatorname{arctg} x$ изображен на рис. 6.15.

Функция $\operatorname{arcctg} x$. Функция $y = \operatorname{ctg} x$ непрерывна в области определения (т. е. при всех $x \in \mathbf{R}$, не равных πn, $n \in \mathbf{Z}$). Выделим на числовой оси Ox промежуток $(0; \pi)$. На этом промежутке функция

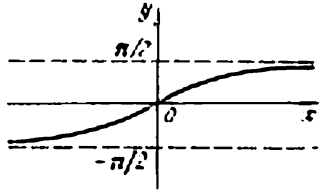


Рис. 6.15

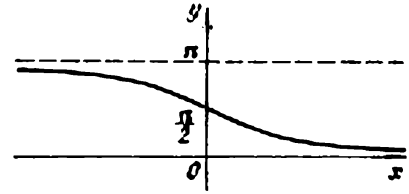


Рис. 6.16

$y = \operatorname{ctg} x$ убывает и принимает все свои значения. Функция $y = \operatorname{ctg} x$, рассматриваемая на промежутке $(0; \pi)$, имеет обратную функцию, которую называют *арккотангенсом* и записывают

$$x = \operatorname{arcctg} y,$$

где y — независимая переменная, а x — зависимая. Обозначая, как обычно, независимую переменную буквой x, а зависимую — буквой y, будем далее писать $y = \operatorname{arcctg} x$.

Свойства функции $\operatorname{arcctg} x$.

1. Область определения — вся числовая прямая.
2. Область изменения (множество значений) — промежуток $(0; \pi)$.
3. Функция $\operatorname{arcctg} x$ не является ни четной, ни нечетной.
4. Функция $\operatorname{arcctg} x$ положительна при всех $x \in \mathbf{R}$.
5. Функция $\operatorname{arcctg} x$ непрерывна и дифференцируема при всех $x \in \mathbf{R}$:

$$(\operatorname{arcctg} x)' = -\frac{1}{1+x^2}.$$

6. Функция $\operatorname{arcctg} x$ убывающая.

График функции $y = \operatorname{arcctg} x$ изображен на рис. 6.16.

1.6. Значения тригонометрических функций некоторых углов. Значения тригонометрических функций углов в 0°, 30°, 45°, 60° и 90° можно вычислить, используя определения соответствующих тригонометрических функций. Значения тригонометрических функций углов в 15°, $\frac{15°}{2}$, $\frac{15°}{4}$, ... можно найти по известным значениям тригонометрических функций 30°, последовательно используя формулы половинного угла (см. пример 1). Значения тригонометрических функций углов, кратных 18°, можно найти, зная значение хотя бы одной тригонометрической функции 18°, например $\sin 18°$ (см. пример 2).

Некоторые значения тригонометрических функций приведены в табл. 2.

Таблица 2

Аргумент	Функция			
	$\sin\alpha$	$\cos\alpha$	$\operatorname{tg}\alpha$	$\operatorname{ctg}\alpha$
$0°\,(0)$	0	1	0	не определен
$15°\left(\frac{\pi}{12}\right)$	$\frac{\sqrt{3}-1}{2\sqrt{2}}$	$\frac{\sqrt{3}+1}{2\sqrt{2}}$	$2-\sqrt{3}$	$2+\sqrt{3}$
$18°\left(\frac{\pi}{10}\right)$	$\frac{\sqrt{5}-1}{4}$	$\frac{\sqrt{5+\sqrt{5}}}{2\sqrt{2}}$	$\frac{\sqrt{5}-1}{\sqrt{10+2\sqrt{5}}}$	$\frac{\sqrt{10+2\sqrt{5}}}{\sqrt{5}-1}$
$30°\left(\frac{\pi}{6}\right)$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}$	$\sqrt{3}$
$36°\left(\frac{\pi}{5}\right)$	$\frac{\sqrt{5-\sqrt{5}}}{2\sqrt{2}}$	$\frac{\sqrt{5}+1}{4}$	$\frac{\sqrt{10-2\sqrt{5}}}{\sqrt{5}+1}$	$\frac{\sqrt{5}+1}{\sqrt{10-2\sqrt{5}}}$
$45°\left(\frac{\pi}{4}\right)$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$		1
$54°\left(\frac{3\pi}{10}\right)$	$\frac{\sqrt{5}+1}{4}$	$\frac{\sqrt{5-\sqrt{5}}}{2\sqrt{2}}$	$\frac{\sqrt{5}+1}{\sqrt{10-2\sqrt{5}}}$	$\frac{\sqrt{10-2\sqrt{5}}}{\sqrt{5}+1}$
$60°\left(\frac{\pi}{3}\right)$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$	$\frac{1}{\sqrt{3}}$
$72°\left(\frac{2\pi}{5}\right)$	$\frac{\sqrt{5+\sqrt{5}}}{2\sqrt{2}}$	$\frac{\sqrt{5}-1}{4}$	$\frac{\sqrt{10+2\sqrt{5}}}{\sqrt{5}-1}$	$\frac{\sqrt{5}-1}{\sqrt{10+2\sqrt{5}}}$
$75°\left(\frac{5\pi}{12}\right)$	$\frac{\sqrt{3}+1}{2\sqrt{2}}$	$\frac{\sqrt{3}-1}{2\sqrt{2}}$	$2+\sqrt{3}$	$2-\sqrt{3}$
$90°\left(\frac{\pi}{2}\right)$	1	0	не определен	0

Пример 1. Вычислить $\cos 15°$.

Косинусы углов α и $\alpha/2$ связаны равенством

$$\cos^2\frac{\alpha}{2}=\frac{1+\cos\alpha}{2};$$

полагая $\alpha=30°$, получаем

$$\cos^2 15°=\frac{1+\cos 30°}{2}=\frac{1+\sqrt{3}/2}{2}=\frac{2+\sqrt{3}}{4},$$

$$\cos 15°=\frac{\sqrt{2+\sqrt{3}}}{2}=\frac{\sqrt{3}+1}{2\sqrt{2}}.$$

Значение $\sin 15°$ можно вычислить, используя тождество

$$\sin^2\alpha+\cos^2\alpha=1,$$

из которого находим

$$\sin 15°=\sqrt{1-\cos^2 15°}=\frac{\sqrt{2-\sqrt{3}}}{2}=\frac{\sqrt{3}-1}{2\sqrt{2}}.$$

Здесь из двух возможных знаков (плюса и минуса) перед радикалом выбран знак плюс, так как 15°—угол, лежащий в первой четверти, и $\sin 15° > 0$.

Пример 2. Вычислить значения тригонометрических функций угла 18°.

Рассмотрим равнобедренный треугольник ABC ($|AB|=|BC|$) с углом при вершине в 36° (рис. 6.17). Углы при основании треугольника ABC равны 72°. Проведем биссектрисы углов BAC и ABC. Отрезки биссектрис, лежащие внутри треугольника, обозначим соответственно через AM и BK.

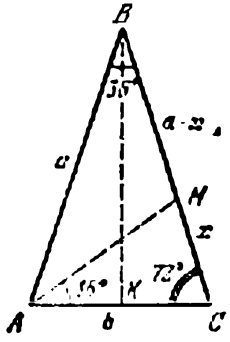


Рис. 6.17

Рассмотрим треугольник AMC. Угол MAC равен 36°, $\angle AMC = \angle MCA$. Следовательно, треугольник AMC подобен треугольнику ABC.

Обозначив $|BC|=a$, $|AC|=b$, $|MC|=x$, запишем условие подобия треугольников ABC и AMC:

$$\frac{|AC|}{|BC|}=\frac{|MC|}{|AC|} \quad \text{или} \quad \frac{b}{a}=\frac{x}{b}. \tag{1}$$

Биссектриса внутреннего угла треугольника делит сторону, которую она пересекает, на отрезки, пропорциональные прилежащим сторонам, поэтому для треугольника ABC имеем

$$\frac{|AB|}{|AC|}=\frac{|MB|}{|MC|}, \quad \text{или} \quad \frac{a}{b}=\frac{a-x}{x}. \tag{2}$$

Рассмотрим прямоугольный треугольник BKC с углом KBC, равным 18°:

$$\frac{|KC|}{|BC|}=\frac{b}{2a}=\sin 18°.$$

Из уравнения (1) получаем $x=b^2/a$. Подставив выражение для x в уравнение (2), получим

$$\frac{a}{b}=\frac{a^2}{b^2}-1.$$

Введем новую переменную $z=b/a$ и запишем последнее уравнение в виде

$$z^2+z-1=0.$$

Корни этого квадратного уравнения

$$z_{1,2}=\frac{-1\pm\sqrt{5}}{2}.$$

По своему геометрическому смыслу величина $b/(2a)$ положительна и равна $(\sqrt{5}-1)/4$. Это число и есть значение $\sin 18°$:

$$\sin 18°=\frac{\sqrt{5}-1}{4}.$$

Далее, используя формулу $\sin^2\alpha+\cos^2\alpha=1$ и определения функции $\operatorname{tg} x$, $\operatorname{ctg} x$, можно найти значения $\cos 18°$, $\operatorname{tg} 18°$, $\operatorname{ctg} 18°$, затем по формулам двойного угла найти $\sin 36°$, $\cos 36°$. Так, $\sin 36°$ можно вычислить следующим образом:

$$\cos 18°=\sqrt{1-\sin^2 18°}=\frac{\sqrt{10+2\sqrt{5}}}{4},$$

$$\sin 36°=2\sin 18°\cos 18°=\frac{1}{2}(\sqrt{5}-1)\sqrt{10+2\sqrt{5}}=\frac{1}{4}\sqrt{10-2\sqrt{5}}.$$

§ 2. Тригонометрические формулы

2.1. Формулы приведения. Вычисление значений тригонометрических функций любого угла сводится к вычислению значений тригонометрических функций острого угла по следующим правилам.

1. Если угол положительный и больший 2π, то функции синус и косинус данного угла приводятся к функциям угла α, $0 < \alpha < 2\pi$, по формулам

$$\left.\begin{aligned}\sin(\alpha+2\pi n) &= \sin\alpha,\\ \cos(\alpha+2\pi n) &= \cos\alpha,\end{aligned}\right\} \quad \alpha\in(0;\,2\pi), \quad n\in\mathbf{Z},$$

а функции тангенс и котангенс данного угла — к функциям угла α, $0 < \alpha < \pi$, по формулам

$$\left.\begin{aligned}\operatorname{tg}(\alpha+\pi n) &= \operatorname{tg}\alpha,\\ \operatorname{ctg}(\alpha+\pi n) &= \operatorname{ctg}\alpha,\end{aligned}\right\} \quad \alpha\in(0;\,\pi), \quad n\in\mathbf{Z}.$$

2. Если угол отрицательный, то тригонометрические функции данного угла приводятся к тригонометрическим функциям положительного угла α по формулам

$$\sin(-\alpha) = -\sin\alpha; \quad \cos(-\alpha) = \cos\alpha;$$
$$\operatorname{tg}(-\alpha) = -\operatorname{tg}\alpha; \quad \operatorname{ctg}(-\alpha) = -\operatorname{ctg}\alpha.$$

3. Тригонометрические функции угла, большего $\pi/2$ и меньшего 2π, приводятся к тригонометрическим функциям острого угла по формулам приведения (см. табл. 3), которые можно сформулировать в виде следующего правила. Если в формуле приведения угол α вычитается из $\pi/2$ или прибавляется к $\pi/2$, взятому нечетное число раз, то приводимая функция меняется на кофункцию *); если же $\pi/2$ взято четное число раз, то название приводимой функции сохраняется. При этом перед приведенной функцией ставится тот знак, который имеет приводимая функция в соответствующей четверти, если считать угол α острым.

Таблица 3

Функция	Аргумент			
	$\beta=\frac{\pi}{2}\pm\alpha$	$\beta=\pi\pm\alpha$	$\beta=\frac{3\pi}{2}\pm\alpha$	$\beta=2\pi-\alpha$
$\sin\beta$	$\cos\alpha$	$\mp\sin\alpha$	$-\cos\alpha$	$-\sin\alpha$
$\cos\beta$	$\mp\sin\alpha$	$-\cos\alpha$	$\pm\sin\alpha$	$\cos\alpha$
$\operatorname{tg}\beta$	$\pm\operatorname{ctg}\alpha$	$\pm\operatorname{tg}\alpha$	$\mp\operatorname{ctg}\alpha$	$-\operatorname{tg}\alpha$
$\operatorname{ctg}\beta$	$\mp\operatorname{tg}\alpha$	$\pm\operatorname{ctg}\alpha$	$\mp\operatorname{tg}\alpha$	$-\operatorname{ctg}\alpha$

2.2. Связь между тригонометрическими функциями одного аргумента. В табл. 4 приведены формулы, связывающие тригонометрические функции одного и того же аргумента. Перед знаком радикала должен быть выбран знак плюс или минус в зависимости от того, в какой четверти находится угол α, а именно так, чтобы знак тригонометрической функции, стоящей в левой части, совпадал со знаком величины, стоящей в правой части равенства.

*) Косинус является *кофункцией* по отношению к синусу, и наоборот. Другая пара кофункций — тангенс и котангенс.

Таблица 4

	$\sin\alpha$	$\cos\alpha$	$\operatorname{tg}\alpha$	$\operatorname{ctg}\alpha$	$\sec\alpha$	$\operatorname{cosec}\alpha$
$\sin\alpha =$		$\pm\sqrt{1-\cos^2\alpha}$	$\dfrac{\operatorname{tg}\alpha}{\pm\sqrt{1+\operatorname{tg}^2\alpha}}$	$\dfrac{1}{\pm\sqrt{1+\operatorname{ctg}^2\alpha}}$	$\dfrac{\pm\sqrt{\sec^2\alpha-1}}{\sec\alpha}$	$\dfrac{1}{\operatorname{cosec}\alpha}$
$\cos\alpha =$	$\pm\sqrt{1-\sin^2\alpha}$		$\dfrac{1}{\pm\sqrt{1+\operatorname{tg}^2\alpha}}$	$\dfrac{\operatorname{ctg}\alpha}{\pm\sqrt{1+\operatorname{ctg}^2\alpha}}$	$\dfrac{1}{\sec\alpha}$	$\dfrac{\pm\sqrt{\operatorname{cosec}^2\alpha-1}}{\operatorname{cosec}\alpha}$
$\operatorname{tg}\alpha =$	$\dfrac{\sin\alpha}{\pm\sqrt{1-\sin^2\alpha}}$	$\dfrac{\pm\sqrt{1-\cos^2\alpha}}{\cos\alpha}$		$\dfrac{1}{\operatorname{ctg}\alpha}$	$\pm\sqrt{\sec^2\alpha-1}$	$\dfrac{1}{\pm\sqrt{\operatorname{cosec}^2\alpha-1}}$
$\operatorname{ctg}\alpha =$	$\dfrac{\pm\sqrt{1-\sin^2\alpha}}{\sin\alpha}$	$\dfrac{\cos\alpha}{\pm\sqrt{1-\cos^2\alpha}}$	$\dfrac{1}{\operatorname{tg}\alpha}$		$\dfrac{1}{\pm\sqrt{\sec^2\alpha-1}}$	$\pm\sqrt{\operatorname{cosec}^2\alpha-1}$
$\sec\alpha =$	$\dfrac{1}{\pm\sqrt{1-\sin^2\alpha}}$	$\dfrac{1}{\cos\alpha}$	$\pm\sqrt{1+\operatorname{tg}^2\alpha}$	$\dfrac{\pm\sqrt{1+\operatorname{ctg}^2\alpha}}{\operatorname{ctg}\alpha}$		$\dfrac{\operatorname{cosec}\alpha}{\pm\sqrt{\operatorname{cosec}^2\alpha-1}}$
$\operatorname{cosec}\alpha =$	$\dfrac{1}{\sin\alpha}$	$\dfrac{1}{\pm\sqrt{1-\cos^2\alpha}}$	$\dfrac{\pm\sqrt{1+\operatorname{tg}^2\alpha}}{\operatorname{tg}\alpha}$	$\pm\sqrt{1+\operatorname{ctg}^2\alpha}$	$\dfrac{\sec\alpha}{\pm\sqrt{\sec^2\alpha-1}}$	

2.3. Тригонометрические функции суммы и разности углов.

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta;$$
$$\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta;$$
$$\operatorname{tg}(\alpha \pm \beta) = \frac{\operatorname{tg}\alpha \pm \operatorname{tg}\beta}{1 \mp \operatorname{tg}\alpha\operatorname{tg}\beta};$$
$$\operatorname{ctg}(\alpha \pm \beta) = \frac{\operatorname{ctg}\alpha\operatorname{ctg}\beta \mp 1}{\operatorname{ctg}\beta \pm \operatorname{ctg}\alpha}.$$

2.4. Тригонометрические функции двойных, тройных и половинных углов.

$$\sin 2\alpha = 2\sin\alpha\cos\alpha;$$
$$\cos 2\alpha = \cos^2\alpha - \sin^2\alpha = 1 - 2\sin^2\alpha = 2\cos^2\alpha - 1;$$
$$\operatorname{tg} 2\alpha = \frac{2\operatorname{tg}\alpha}{1 - \operatorname{tg}^2\alpha}; \qquad \operatorname{ctg} 2\alpha = \frac{\operatorname{ctg}^2\alpha - 1}{2\operatorname{ctg}\alpha};$$
$$\sin 3\alpha = 3\sin\alpha - 4\sin^3\alpha; \qquad \cos 3\alpha = 4\cos^3\alpha - 3\cos\alpha;$$
$$\operatorname{tg} 3\alpha = \frac{3\operatorname{tg}\alpha - \operatorname{tg}^3\alpha}{1 - 3\operatorname{tg}^2\alpha}; \qquad \operatorname{ctg} 3\alpha = \frac{\operatorname{ctg}^3\alpha - 3\operatorname{ctg}\alpha}{3\operatorname{ctg}^2\alpha - 1};$$
$$\sin\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{2}}; \qquad \cos\frac{\alpha}{2} = \pm\sqrt{\frac{1+\cos\alpha}{2}};$$
$$\operatorname{tg}\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{1+\cos\alpha}} = \frac{\sin\alpha}{1+\cos\alpha} = \frac{1-\cos\alpha}{\sin\alpha};$$
$$\operatorname{ctg}\frac{\alpha}{2} = \pm\sqrt{\frac{1+\cos\alpha}{1-\cos\alpha}} = \frac{\sin\alpha}{1-\cos\alpha} = \frac{1+\cos\alpha}{\sin\alpha}.$$

В формулах половинного угла знак перед радикалом берется в зависимости от знака тригонометрической функции, стоящей в левой части равенства.

2.5. Преобразование суммы (разности) тригонометрических функций в произведение (преобразование тригонометрических выражений к виду, удобному для логарифмирования).

$$\sin\alpha + \sin\beta = 2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2};$$
$$\sin\alpha - \sin\beta = 2\cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2};$$
$$\cos\alpha + \cos\beta = 2\cos\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2};$$
$$\cos\alpha - \cos\beta = -2\sin\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2} = 2\sin\frac{\alpha+\beta}{2}\sin\frac{\beta-\alpha}{2};$$
$$\cos\alpha + \sin\alpha = \sqrt{2}\cos(45^\circ - \alpha); \qquad \cos\alpha - \sin\alpha = \sqrt{2}\sin(45^\circ - \alpha);$$
$$\operatorname{tg}\alpha \pm \operatorname{tg}\beta = \frac{\sin(\alpha \pm \beta)}{\cos\alpha\cos\beta}; \qquad \operatorname{ctg}\alpha \pm \operatorname{ctg}\beta = \frac{\sin(\beta \pm \alpha)}{\sin\alpha\sin\beta};$$
$$\operatorname{tg}\alpha + \operatorname{ctg}\beta = \frac{\cos(\alpha-\beta)}{\cos\alpha\sin\beta}; \qquad \operatorname{tg}\alpha - \operatorname{ctg}\beta = -\frac{\cos(\alpha+\beta)}{\cos\alpha\sin\beta};$$
$$\operatorname{tg}\alpha + \operatorname{ctg}\alpha = 2\operatorname{cosec} 2\alpha; \qquad \operatorname{tg}\alpha - \operatorname{ctg}\alpha = -2\operatorname{ctg} 2\alpha;$$
$$1 + \cos\alpha = 2\cos^2\frac{\alpha}{2}; \qquad 1 - \cos\alpha = 2\sin^2\frac{\alpha}{2};$$

$$1+\sin\alpha=2\cos^2\left(45^\circ-\frac{\alpha}{2}\right);\qquad 1-\sin\alpha=2\sin^2\left(45^\circ-\frac{\alpha}{2}\right);$$

$$1\pm\operatorname{tg}\alpha=\frac{\sin(45^\circ\pm\alpha)}{\cos 45^\circ\cos\alpha}=\frac{\sqrt{2}\sin(45^\circ\pm\alpha)}{\cos\alpha};$$

$$1\pm\operatorname{tg}\alpha\operatorname{tg}\beta=\frac{\cos(\alpha\mp\beta)}{\cos\alpha\cos\beta};\quad \operatorname{ctg}\alpha\operatorname{ctg}\beta\pm 1=\frac{\cos(\alpha\mp\beta)}{\sin\alpha\sin\beta};$$

$$1-\operatorname{tg}^2\alpha=\frac{\cos 2\alpha}{\cos^2\alpha};\qquad 1-\operatorname{ctg}^2\alpha=-\frac{\cos 2\alpha}{\sin^2\alpha};$$

$$\operatorname{tg}^2\alpha-\operatorname{tg}^2\beta=\frac{\sin(\alpha+\beta)\sin(\alpha-\beta)}{\cos^2\alpha\cos^2\beta};$$

$$\operatorname{ctg}^2\alpha-\operatorname{ctg}^2\beta=\frac{\sin(\alpha+\beta)\sin(\beta-\alpha)}{\sin^2\alpha\sin^2\beta};$$

$$\operatorname{tg}^2\alpha-\sin^2\alpha=\operatorname{tg}^2\alpha\sin^2\alpha;\quad \operatorname{ctg}^2\alpha-\cos^2\alpha=\operatorname{ctg}^2\alpha\cos^2\alpha.$$

2.6. Преобразование произведения тригонометрических функций в сумму.

$$\sin\alpha\sin\beta=\frac{1}{2}[\cos(\alpha-\beta)-\cos(\alpha+\beta)];$$

$$\cos\alpha\cos\beta=\frac{1}{2}[\cos(\alpha-\beta)+\cos(\alpha+\beta)];$$

$$\sin\alpha\cos\beta=\frac{1}{2}[\sin(\alpha-\beta)+\sin(\alpha+\beta)];$$

$$\sin\alpha\sin\beta\sin\gamma=\frac{1}{4}[\sin(\alpha+\beta-\gamma)+\sin(\beta+\gamma-\alpha)+$$
$$+\sin(\gamma+\alpha-\beta)-\sin(\alpha+\beta+\gamma)];$$

$$\sin\alpha\cos\beta\cos\gamma=\frac{1}{4}[\sin(\alpha+\beta-\gamma)-\sin(\beta+\gamma-\alpha)+$$
$$+\sin(\gamma+\alpha-\beta)+\sin(\alpha+\beta+\gamma)];$$

$$\sin\alpha\sin\beta\cos\gamma=$$
$$=\frac{1}{4}[-\cos(\alpha+\beta-\gamma)+\cos(\beta+\gamma-\alpha)+\cos(\gamma+\alpha-\beta)-\cos(\alpha+\beta+\gamma)];$$

$$\cos\alpha\cos\beta\cos\gamma=$$
$$=\frac{1}{4}[\cos(\alpha+\beta-\gamma)+\cos(\beta+\gamma-\alpha)+\cos(\gamma+\alpha-\beta)+\cos(\alpha+\beta+\gamma)].$$

2.7. Простейшие соотношения между обратными тригонометрическими функциями.

$$\operatorname{arcsin}\alpha=-\operatorname{arcsin}(-\alpha)=\frac{\pi}{2}-\arccos\alpha=\operatorname{arctg}\frac{\alpha}{\sqrt{1-\alpha^2}};$$

$$\arccos\alpha=\pi-\arccos(-\alpha)=\frac{\pi}{2}-\operatorname{arcsin}\alpha=\operatorname{arcctg}\frac{\alpha}{\sqrt{1-\alpha^2}};$$

$$\operatorname{arctg}\alpha=-\operatorname{arctg}(-\alpha)=\frac{\pi}{2}-\operatorname{arcctg}\alpha=\operatorname{arcsin}\frac{\alpha}{\sqrt{1+\alpha^2}};$$

$$\operatorname{arcctg}\alpha=\pi-\operatorname{arcctg}(-\alpha)=\frac{\pi}{2}-\operatorname{arctg}\alpha=\arccos\frac{\alpha}{\sqrt{1+\alpha^2}}.$$

Более сложная зависимость между обратными тригонометрическими функциями может быть установлена методами, используемыми при решении уравнений, содержащих обратные тригонометрические функции (см. п. 3.4).

Пример. Выразить сумму двух обратных тригонометрических функций $\arcsin x+\arcsin y$ через одну обратную тригонометрическую функцию (например, арксинус).

Обозначим данную сумму буквой z:

$$\arcsin x+\arcsin y=z. \tag{1}$$

Это равенство можно рассматривать как уравнение с тремя неизвестными x, y и z. Как следствие уравнения (1) получаем уравнение

$$\sin(\arcsin x+\arcsin y)=\sin z,$$

которое приводится к виду

$$\sin(\arcsin x)\cdot\cos(\arcsin y)+\sin(\arcsin y)\cdot\cos(\arcsin x)=$$

$$=\sin z \Leftrightarrow x\sqrt{1-y^2}+y\sqrt{1-x^2}=\sin z.$$

Теперь нужно решить уравнение относительно переменной z, учитывая, что $z\in[-\pi;\ \pi]$.

Множество решений этого уравнения будет иметь вид

$$z=\arcsin\left(x\sqrt{1-y^2}+y\sqrt{1-x^2}\right)$$

при $xy\leqslant 0$ или $x^2+y^2\leqslant 1$;

$$z=\pi-\arcsin\left(x\sqrt{1-y^2}+y\sqrt{1-x^2}\right)$$

при $x>0$, $y>0$ и $x^2+y^2>1$;

$$z=-\pi-\arcsin\left(x\sqrt{1-y^2}+y\sqrt{1-x^2}\right)$$

при $x<0$, $y<0$ и $x^2+y^2>1$.

Полученные выражения для переменной z и дают искомое выражение суммы двух обратных тригонометрических функций через одну обратную тригонометрическую функцию.

§ 3. Решение тригонометрических уравнений и неравенств

3.1. Простейшие тригонометрические уравнения.

Решение уравнения $\sin x=a$. Рассмотрим функцию $y=\sin x$ на промежутке $[-\pi/2;\ \pi/2]$. На этом промежутке функция $y=\sin x$ возрастает, принимая все свои значения — от наименьшего,

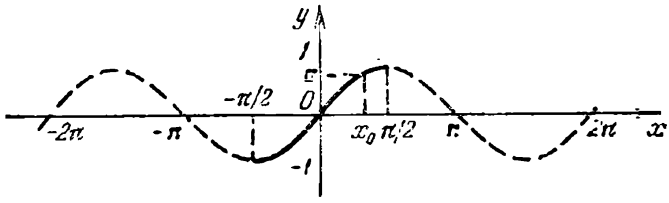


Рис. 6.18

равного -1, на левом конце промежутка до наибольшего, равного 1, на правом конце промежутка (рис. 6.18). В силу непрерывности функции $y=\sin x$ каждому значению $y=a$, удовлетворяющему условию $|a|\leqslant 1$, соответствует единственное значение $x_0\in[-\pi/2;\ \pi/2]$

такое, что

$$\sin x_0 = a.$$

Угол x_0 есть арксинус числа a (обозначается $\arcsin a$).

Уравнение вида $\sin x = a$ при $|a| \leqslant 1$ имеет множество решений $x = (-1)^n \arcsin a + \pi n$, $n \in \mathbf{Z}$; при $|a| > 1$ уравнение решений не имеет (т. е. множество решений — пустое множество).

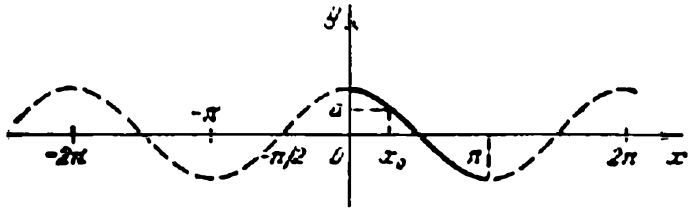


Рис. 6.19

Решение уравнения $\cos x = a$. Рассмотрим функцию $y = \cos x$ на промежутке $[0; \pi]$. На этом промежутке функция $y = \cos x$ убывает, принимая все свои значения — от наибольшего, равного 1, на левом конце промежутка до наименьшего, равного -1, на правом конце промежутка (рис. 6.19). В силу непрерывности функции $y = \cos x$ каждому значению $y = a$, удовлетворяющему условию $|a| \leqslant 1$, соответствует единственное значение $x_0 \in [0; \pi]$ такое, что

$$\cos x_0 = a.$$

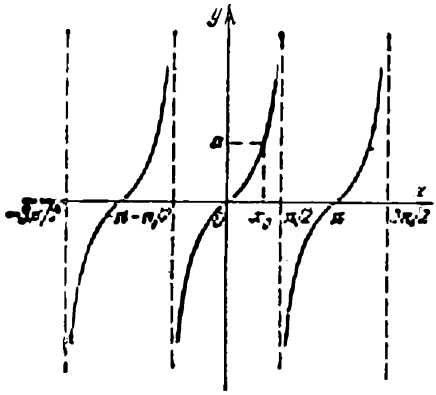


Рис. 6.20

Угол x_0 есть арккосинус числа a (обозначается $\arccos a$).

Уравнение вида $\cos x = a$ при $|a| \leqslant 1$ имеет множество решений $x = \pm \arccos a + 2\pi n$, $n \in \mathbf{Z}$; при $|a| > 1$ уравнение решений не имеет (множество решений — пустое множество).

Решение уравнения $\operatorname{tg} x = a$. Функция $y = \operatorname{tg} x$ на промежутке $(-\pi/2; \pi/2)$ возрастает и принимает все свои значения (рис. 6.20). В силу непрерывности функции $y = \operatorname{tg} x$ каждому значению $y = a$ соответствует единственное значение $x_0 \in (-\pi/2; \pi/2)$ такое, что

$$\operatorname{tg} x_0 = a.$$

Угол x_0 есть арктангенс числа a (обозначается $\operatorname{arctg} a$).

Уравнение вида $\operatorname{tg} x = a$ при любом a имеет множество решений

$$a = \operatorname{arctg} a + \pi n, \quad n \in \mathbf{Z}.$$

Решение уравнения $\operatorname{ctg} x = a$. Функция $y = \operatorname{ctg} x$ на промежутке $(0; \pi)$ убывает и принимает все свои значения (рис. 6.21). В силу непрерывности функции $y = \operatorname{ctg} x$ каждому значению $y = a$

соответствует единственное значение $x_0 \in (0; \pi)$ такое, что

$$\operatorname{ctg} x_0 = a.$$

Угол x_0 есть арккотангенс числа a (обозначается $\operatorname{arcctg} a$).

Уравнение вида $\operatorname{ctg} x = a$ при любом a имеет множество решений $x = \operatorname{arcctg} a + \pi n$, $n \in Z$.

3.2. Примеры более сложных тригонометрических уравнений.

1. Тригонометрическое уравнение вида

$$P(z, u, v, w) = 0, \qquad (1)$$

где P — многочлен аргументов $z = \sin kx$, $u = \cos nx$, $v = \operatorname{tg} mx$, $w = \operatorname{ctg} lx$ (k, n, m, l — натуральные числа), с помощью формул для тригонометрических функций суммы углов (в частности, формул двойного и тройного угла) можно свести к рациональному уравнению относительно аргументов $\sin x$, $\cos x$, $\operatorname{tg} x$ и $\operatorname{ctg} x$. После этого уравнение (1) может быть сведено к рациональному уравнению относительно неизвестной $t = \operatorname{tg} \frac{x}{2}$ с помощью формул *универсальной подстановки*

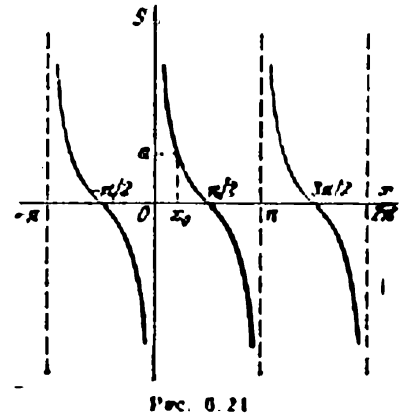


Рис. 6.21

$$\sin x = \frac{2 \operatorname{tg} \frac{x}{2}}{1 + \operatorname{tg}^2 \frac{x}{2}}; \quad \cos x = \frac{1 - \operatorname{tg}^2 \frac{x}{2}}{1 + \operatorname{tg}^2 \frac{x}{2}};$$

$$\operatorname{tg} x = \frac{2 \operatorname{tg} \frac{x}{2}}{1 - \operatorname{tg}^2 \frac{x}{2}}; \quad \operatorname{ctg} x = \frac{1 - \operatorname{tg}^2 \frac{x}{2}}{2 \operatorname{tg} \frac{x}{2}}.$$

Пример 1. Решить уравнение

$$\sin 2x + \operatorname{tg} x = 2.$$

Выражая $\sin 2x$ через $\operatorname{tg} x$, получаем уравнение

$$\frac{2 \operatorname{tg} x}{1 + \operatorname{tg}^2 x} + \operatorname{tg} x = 2,$$

которое приводится к кубическому уравнению относительно неизвестной $t = \operatorname{tg} x$:

$$t^3 - 2t^2 + 3t - 2 = 0 \Leftrightarrow (t-1)(t^2 - t + 2) = 0.$$

Легко убедиться, что данное уравнение имеет только один действительный корень: $t = 1$.

Решая уравнение $\operatorname{tg} x = 1$, находим множество решений исходного уравнения: $x = \frac{\pi}{4} + \pi k$, $k \in Z$.

Общий метод решения тригонометрического уравнения с помощью формул универсальной тригонометрической подстановки не всегда

удобен, так как в ряде случаев он может свести решение к нахождению корней многочлена довольно высокой степени.

2. Уравнение вида

$$P(\sin x+\cos x,\ \sin x\cos x)=0, \tag{2}$$

где P — многочлен аргументов $t=\sin x+\cos x$, $z=\sin x\cos x$, может быть сведено к уравнению относительно неизвестной $t=\sin x+\cos x$, если воспользоваться тригонометрическим тождеством

$$(\sin x+\cos x)^2=\sin^2 x+\cos^2 x+2\sin x\cos x=1+2\sin x\cos x,$$

из которого следует, что

$$\sin x\cos x=\frac{t^2-1}{2}. \tag{3}$$

Теперь можно привести уравнение (2) к виду

$$P\left(t,\ \frac{t^2-1}{2}\right)=0.$$

Аналогично уравнение вида

$$P(\sin x-\cos x,\ \sin x\cos x)=0$$

заменой $\sin x-\cos x=t$ сводится к уравнению

$$R\left(t,\ \frac{1-t^2}{2}\right)=0.$$

Пример 2. Решить уравнение

$$\sin x+\cos x-2\sqrt{2}\sin x\cos x=0.$$

Обозначая $\sin x+\cos x=t$ и пользуясь равенством

$$\sin x\cos x=\frac{t^2-1}{2},$$

сводим уравнение к квадратному уравнению относительно t:

$$\sqrt{2}\,t^2-t-\sqrt{2}=0;$$

корнями этого уравнения будут числа

$$t_1=\sqrt{2}\quad\text{и}\quad t_2=-1/\sqrt{2}.$$

Таким образом, решение исходного уравнения сводится к решению двух более простых тригонометрических уравнений:

$$\sin x+\cos x=\sqrt{2}\quad\text{и}\quad\sin x+\cos x=-\frac{1}{\sqrt{2}}.$$

Умножая обе части полученных уравнений на $1/\sqrt{2}$, сводим уравнения к двум более простым тригонометрическим уравнениям:

$$\frac{1}{\sqrt{2}}\sin x+\frac{1}{\sqrt{2}}\cos x=1\Leftrightarrow\sin x\cos\frac{\pi}{4}+\sin\frac{\pi}{4}\cos x=1\Leftrightarrow\sin\left(x+\frac{\pi}{4}\right)=1,$$

$$\frac{1}{\sqrt{2}}\sin x+\frac{1}{\sqrt{2}}\cos x=-\frac{1}{2}\Leftrightarrow\sin\left(x+\frac{\pi}{4}\right)=-\frac{1}{2}.$$

Множества решений уравнений $\sin\left(x+\frac{\pi}{4}\right)=1$ и $\sin\left(x+\frac{\pi}{4}\right)=-\frac{1}{2}$ будут соответственно иметь вид

$$x=\frac{\pi}{4}+2\pi k,\qquad k\in Z,$$

$$x=(-1)^{n+1}\frac{\pi}{6}-\frac{\pi}{4}+\pi n,\qquad n\in Z.$$

3. Уравнение вида $a\sin x+b\cos x=c$ (a, b, c — действительные числа) может быть решено при помощи формул универсальной подстановки (см. 1).

Кроме того, это уравнение может быть решено методом *дополнительного угла*.

Разделим обе части исходного уравнения на $\sqrt{a^2+b^2}\neq 0$ (если $a=b=0$, то уравнение либо превращается в тождество (при $c=0$), либо не имеет решений (при $c\neq 0$)). После деления получаем уравнение, эквивалентное исходному:

$$\frac{a}{\sqrt{a^2+b^2}}\sin x+\frac{b}{\sqrt{a^2+b^2}}\cos x=\frac{c}{\sqrt{a^2+b^2}}. \qquad (4)$$

Легко проверить, что коэффициенты $a/\sqrt{a^2+b^2}$ и $b/\sqrt{a^2+b^2}$ связаны равенством

$$\left(\frac{a}{\sqrt{a^2+b^2}}\right)^2+\left(\frac{b}{\sqrt{a^2+b^2}}\right)^2=1,$$

а потому их можно считать значениями синуса и косинуса некоторого угла φ:

$$\frac{a}{\sqrt{a^2+b^2}}=\sin\varphi, \quad \frac{b}{\sqrt{a^2+b^2}}=\cos\varphi.$$

Таким образом, уравнение (4) приводится к уравнению

$$\cos x\cos\varphi+\sin x\sin\varphi=\frac{c}{\sqrt{a^2+b^2}} \Leftrightarrow \cos(x-\varphi)=\frac{c}{\sqrt{a^2+b^2}},$$

множество решений которого при $|c|\leqslant\sqrt{a^2+b^2}$ имеет вид

$$x=\varphi\pm\arccos\frac{c}{\sqrt{a^2+b^2}}+2\pi n, \quad n\in\mathbf{Z}.$$

Осталось найти какое-нибудь значение угла φ, являющееся решением системы тригонометрических уравнений

$$\sin\varphi=\frac{a}{\sqrt{a^2+b^2}}, \quad \cos\varphi=\frac{b}{\sqrt{a^2+b^2}}.$$

Здесь возможны следующие случаи (для φ выбраны наиболее простые из нескольких возможных выражений):

если $a>0$, $b>0$, то

$$\varphi=\arcsin\frac{a}{\sqrt{a^2+b^2}}=\arccos\frac{b}{\sqrt{a^2+b^2}}=\operatorname{arctg}\frac{a}{b};$$

если $a>0$, $b<0$, то $\varphi=\arccos\dfrac{b}{\sqrt{a^2+b^2}}$;

если $a<0$, $b<0$, то $\varphi=\pi+\operatorname{arctg}\dfrac{a}{b}$;

если $a<0$, $b>0$, то $\varphi=\arcsin\dfrac{a}{\sqrt{a^2+b^2}}=\operatorname{arctg}\dfrac{a}{b}$.

4. Уравнение вида

$$\sin^p x+\cos^k x=1, \quad p, k=3, 4, 5, \ldots,$$

может быть решено следующим способом.

Из определения функций $\sin x$ и $\cos x$ следует, что при всех значениях аргумента $x \neq \pi n/2$, $n \in \mathbf{Z}$, справедливы неравенства

$$\sin^p x \leqslant |\sin x|^p < \sin^2 x, \quad \cos^k x \leqslant |\cos x|^k < \cos^2 x.$$

Как следствие этих неравенств получаем неравенство

$$\sin^p x + \cos^k x \leqslant |\sin x|^p + |\cos x|^k < 1.$$

Следовательно, решениями исходного уравнения могут быть лишь числа из множества значений аргумента $x = \pi n/2$, $n \in \mathbf{Z}$.

Далее процедура нахождения корней сводится к проверке — какие из указанных значений аргумента являются корнями данного уравнения.

Аналогично решаются уравнения вида

$$|\sin x|^\alpha + |\cos x|^\beta = 1,$$

где α, $\beta < 2$.

3.3. Решение простейших тригонометрических неравенств. В табл. 5 приведены множества решений простейших тригонометрических неравенств. Более сложные неравенства решаются методами, сходными с методами решений тригонометрических уравнений.

Таблица 5

Вид неравенства	Множество решений неравенства ($n \in \mathbf{Z}$)
$\sin x > a \quad (\lvert a \rvert < 1)$	$x \in (\arcsin a + 2\pi n;\ \pi - \arcsin a + 2\pi n)$
$\sin x < a \quad (\lvert a \rvert < 1)$	$x \in (-\pi - \arcsin a + 2\pi n;\ \arcsin a + 2\pi n)$
$\cos x > a \quad (\lvert a \rvert < 1)$	$x \in (-\arccos a + 2\pi n;\ \arccos a + 2\pi n)$
$\cos x < a \quad (\lvert a \rvert < 1)$	$x \in (\arccos a + 2\pi n;\ 2\pi - \arccos a + 2\pi n)$
$\operatorname{tg} x > a$	$x \in \left(\operatorname{arctg} a + \pi n;\ \frac{\pi}{2} + \pi n\right)$
$\operatorname{tg} x < a$	$x \in \left(-\frac{\pi}{2} + \pi n;\ \operatorname{arctg} a + \pi n\right)$
$\operatorname{ctg} x > a$	$x \in (\pi n;\ \operatorname{arcctg} a + \pi n)$
$\operatorname{ctg} x < a$	$x \in (\operatorname{arcctg} a + \pi n;\ \pi + \pi n)$

3.4. Примеры решения уравнений и неравенств, содержащих обратные тригонометрические функции. В табл. 6 приведены мно-

Таблица 6

Вид неравенства	Множество решений неравенства
$\arcsin x > a \quad (\lvert a \rvert < \pi/2)$	$x \in (\sin a;\ 1]$
$\arcsin x < a \quad (\lvert a \rvert \leqslant \pi/2)$	$x \in [-1;\ \sin a)$
$\arccos x > a \quad (0 < a < \pi)$	$x \in [-1;\ \cos a)$
$\arccos x < a \quad (0 < a \leqslant \pi)$	$x \in (\cos a;\ 1]$
$\operatorname{arctg} x > a \quad (\lvert a \rvert < \pi/2)$	$x \in (\operatorname{tg} a;\ +\infty)$
$\operatorname{arctg} x < a \quad (\lvert a \rvert < \pi/2)$	$x \in (-\infty;\ \operatorname{tg} a)$
$\operatorname{arcctg} x > a \quad (0 < a < \pi)$	$x \in (-\infty;\ \operatorname{ctg} a)$
$\operatorname{arcctg} x < a \quad (0 < a < \pi)$	$x \in (\operatorname{ctg} a;\ +\infty)$

жества решений простейших неравенств для обратных тригонометрических функций.

Рассмотрим несколько примеров решения более сложных уравнений и неравенств, содержащих обратные тригонометрические функции.

Пример 1. Решить уравнение

$$\arcsin x+\arcsin 2x=\frac{\pi}{3}. \tag{5}$$

Множество допустимых значений неизвестного x — промежуток $[-1/2;\ 1/2]$. Обозначая $\arcsin 2x=\alpha$, $\arcsin x=\beta$, получим

$$\begin{aligned} &\sin\alpha=2x, \quad \alpha\in[-\pi/2;\ \pi/2];\\ &\sin\beta=x, \quad \beta\in[-\pi/6;\ \pi/6]. \end{aligned} \tag{6}$$

В новых обозначениях уравнение (5) запишется в виде

$$\alpha+\beta=\frac{\pi}{3}\Leftrightarrow\alpha=\frac{\pi}{3}-\beta. \tag{7}$$

Из условий (6) следует, что углы α и $\pi/3-\beta$ принадлежат промежутку $[-\pi/2;\ \pi/2]$. Следовательно, уравнение (7) эквивалентно уравнению

$$\sin\alpha=\sin\left(\frac{\pi}{3}-\beta\right)\Leftrightarrow 2x=\sin\frac{\pi}{3}\cos\beta-\sin\beta\cos\frac{\pi}{3}. \tag{8}$$

Из равенства $\sin^2\beta+\cos^2\beta=1$ с учетом условий (6) следует, что

$$\cos\beta=\sqrt{1-x^2}.$$

Следовательно, уравнение (8) можно записать в виде

$$2x=\frac{\sqrt{3}}{2}\sqrt{1-x^2}-\frac{x}{2}\Leftrightarrow 4x=\sqrt{3}\sqrt{1-x^2}-x\Leftrightarrow 5x=\sqrt{3}\sqrt{1-x^2}.$$

Возведя обе части последнего уравнения в квадрат, находим единственное решение этого иррационального уравнения:

$$x=\frac{1}{2}\sqrt{\frac{3}{7}}.$$

Найденное значение x и есть решение исходного тригонометрического уравнения (5).

Пример 2. Найти область определения функции

$$f(x)=\arcsin(\arcsin x)+\arccos\left(\frac{2\arccos x}{\pi-2}\right).$$

Область определения данной функции находится как множество решений следующей системы неравенств:

$$\begin{aligned} &-1\leqslant x\leqslant 1,\\ &-1\leqslant\arcsin x\leqslant 1,\\ &-1\leqslant\frac{2\arccos x}{\pi-2}\leqslant 1. \end{aligned}$$

Решим второе неравенство системы:

$$-1\leqslant\arcsin x\leqslant 1. \tag{9}$$

Так как множество значений функции арксинус — промежуток $[-\pi/2,\ \pi/2]$, то все углы, фигурирующие в неравенстве (9) (-1, $\arcsin x$ и 1), принадлежат одному и тому же промежутку монотонного возрастания функции синус. Следовательно, неравенство (9) эквивалентно неравенству

$$\sin(-1)\leqslant\sin(\arcsin x)\leqslant\sin 1\Leftrightarrow-\sin 1\leqslant x\leqslant\sin 1.$$

Решим третье неравенство системы:

$$-1\leqslant\frac{2\arccos x}{\pi-2}\leqslant 1. \tag{10}$$

Умножая все три части двойного неравенства на положительную величину $(\pi-2)/2$, получим неравенство, эквивалентное неравенству (10):

$$-\frac{\pi-2}{2} \leqslant \arccos x \leqslant \frac{\pi-2}{2}.$$

Это двойное неравенство эквивалентно системе двух неравенств

$$\arccos x \geqslant -\frac{\pi-2}{2},$$

$$\arccos x \leqslant \frac{\pi-2}{2}.$$

причем первое неравенство справедливо при всех значениях $x \in [-1; 1]$, так как множество значений функции арккосинус — промежуток $[0; \pi]$.

Решим второе неравенство системы. Так как величины, стоящие в обеих частях неравенства, принадлежат промежутку монотонного убывания функции $\cos x$, то это неравенство эквивалентно неравенству

$$\cos(\arccos x) \geqslant \cos\left(\frac{\pi}{2}-1\right) = \sin 1.$$

Таким образом, область определения данной функции находится как множество значений x, удовлетворяющих системе

$$\begin{gathered} -1 \leqslant x \leqslant 1, \\ -\sin 1 \leqslant x \leqslant \sin 1, \\ x \geqslant \sin 1. \end{gathered}$$

Эта система имеет единственное решение $x=\sin 1$, и, следовательно, область определения данной функции $f(x)$ состоит из одной точки $x=\sin 1$.

Пример 3. Решить уравнение

$$2\operatorname{arctg} x + \arcsin\frac{2x}{1+x^2} = \pi. \tag{11}$$

Множество допустимых значений неизвестной определяется неравенством

$$-1 \leqslant \frac{2x}{1+x^2} \leqslant 1,$$

которое выполняется при любом $x \in \mathbf{R}$.

Обозначив $\operatorname{arctg} x = \alpha$, $\arcsin\frac{2x}{1+x^2} = \beta$, запишем уравнение (11) в виде

$$2\alpha = \pi - \beta. \tag{12}$$

Вычислив функцию синус от обеих частей уравнения, получим уравнение, являющееся следствием уравнения (12):

$$\sin 2\alpha = \sin\beta,$$

или

$$\sin 2\alpha = \frac{2x}{1+x^2}. \tag{13}$$

Согласно принятым выше обозначениям $\operatorname{tg}\alpha = x$, $\alpha \in \left(-\frac{\pi}{2}; \frac{\pi}{2}\right)$, поэтому на основании формулы

$$1+\operatorname{tg}^2\alpha = \frac{1}{\cos^2\alpha}$$

находим

$$\cos^2\alpha = \frac{1}{1+x^2},$$

откуда следует, что

$$\cos\alpha = \frac{1}{\sqrt{1+x^2}} \tag{14}$$

(при получении последнего равенства учтено, что $\alpha \in \left(-\frac{\pi}{2}; \frac{\pi}{2}\right)$.)

Далее, как при $x \geqslant 0$, так и при $x < 0$ из тождества $\sin^2\alpha + \cos^2\alpha = 1$ находим

$$\sin\alpha = \frac{x}{\sqrt{1+x^2}}. \tag{15}$$

По формуле двойного угла $\sin 2\alpha = 2\sin\alpha\cos\alpha$, используя равенства (14), (15), находим

$$\sin 2\alpha = \frac{2x}{1+x^2}.$$

Следовательно, при $\alpha \in \left(-\frac{\pi}{2}; \frac{\pi}{2}\right)$ уравнение (13) обращается в тождество

$$\frac{2x}{1+x^2} = \frac{2x}{1+x^2}.$$

Таким образом, оказалось, что уравнение (13), являющееся следствием уравнения (12), представляет собой тождество. Для того чтобы выяснить вопрос о том, при каких значениях x уравнение (12) обращается в тождество, необходимо установить, при каких значениях x уравнение (12) эквивалентно уравнению (13).

Уравнение (13) эквивалентно уравнению (12) лишь для тех значений x, при которых левая и правая части уравнения (12) принадлежат одному и тому же промежутку монотонности функции синус.

Из определения функции арктангенс и арксинус имеем

$$-\frac{\pi}{2} < \alpha < \frac{\pi}{2} \Leftrightarrow -\pi < 2\alpha < \pi,$$

$$-\frac{\pi}{2} \leqslant \beta \leqslant \frac{\pi}{2} \Leftrightarrow -\frac{\pi}{2} \leqslant -\beta \leqslant \frac{\pi}{2} \Leftrightarrow \frac{\pi}{2} \leqslant \pi - \beta \leqslant \frac{3\pi}{2}.$$

Из приведенных неравенств следует, что исходное уравнение может обратиться в тождество лишь при тех значениях x, для которых левая и правая части уравнения (12) соответственно удовлетворяют неравенствам

$$\pi/2 \leqslant 2\alpha < \pi, \quad \pi/2 \leqslant \pi - \beta < \pi,$$

из которых получаем, что

$$\pi/4 \leqslant \alpha < \pi/2, \quad 0 < \beta \leqslant \pi/2,$$

т. е. значения x должны удовлетворять неравенствам

$$\frac{\pi}{4} \leqslant \operatorname{arctg} x < \frac{\pi}{2}, \quad 0 < \arcsin\frac{2x}{1+x^2} \leqslant \frac{\pi}{2}.$$

Множество решений второго неравенства: $x > 0$; множество решений первого: $x \geqslant 1$.

Следовательно, уравнения (12) и (13) эквивалентны при $x \in [1; +\infty)$, а так как уравнение (13) обращается в тождество, то и уравнение (12) обращается в тождество при $x \in [1; +\infty)$. Эти значения x и являются решениями данного уравнения (11).

§ 4. Соотношения между элементами треугольника

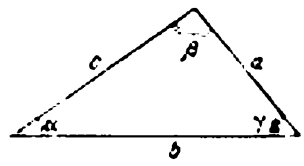


Рис. 6.22

4.1. Основные формулы. Обозначения: a, b, c — стороны треугольника; α, β, γ — углы треугольника (рис. 6.22); $p = \frac{a+b+c}{2}$ — полупериметр; S — площадь; R — радиус описанной окружности; r — радиус вписанной окружности; h — высота; m — медиана; l — биссектриса; индексы a, b, c при h, m, l конкретизируют одну из трех высот, медиан и биссектрис соответственно (например, m_a — медиана, проведенная к

стороне a); r_a — радиус вневписанной окружности, касающейся стороны a и продолжения сторон b и c.

1. Теорема косинусов. Квадрат стороны треугольника равен сумме квадратов двух других сторон без удвоенного произведения этих сторон на косинус угла между ними:

$$a^2 = b^2 + c^2 - 2bc\cos\alpha;$$
$$b^2 = a^2 + c^2 - 2ac\cos\beta;$$
$$c^2 = a^2 + b^2 - 2ab\cos\gamma.$$

2. Теорема синусов. Стороны треугольника пропорциональны синусам противолежащих углов:

$$\frac{a}{\sin\alpha} = \frac{b}{\sin\beta} = \frac{c}{\sin\gamma} = 2R.$$

3. Теорема тангенсов.

$$\frac{a+b}{a-b} = \frac{\operatorname{tg}\frac{\alpha+\beta}{2}}{\operatorname{tg}\frac{\alpha-\beta}{2}} = \frac{\operatorname{ctg}\frac{\gamma}{2}}{\operatorname{tg}\frac{\alpha-\beta}{2}}; \quad \frac{a+c}{a-c} = \frac{\operatorname{tg}\frac{\alpha+\gamma}{2}}{\operatorname{tg}\frac{\alpha-\gamma}{2}} = \frac{\operatorname{ctg}\frac{\beta}{2}}{\operatorname{tg}\frac{\alpha-\gamma}{2}};$$

$$\frac{b+c}{b-c} = \frac{\operatorname{tg}\frac{\beta+\gamma}{2}}{\operatorname{tg}\frac{\beta-\gamma}{2}} = \frac{\operatorname{ctg}\frac{\alpha}{2}}{\operatorname{tg}\frac{\beta-\gamma}{2}}.$$

4. Формулы вычисления площади треугольника.

$$S = \frac{1}{2}ab\sin\gamma; \quad S = \sqrt{p(p-a)(p-b)(p-c)};$$

$$S = p^2\operatorname{tg}\frac{\alpha}{2}\operatorname{tg}\frac{\beta}{2}\operatorname{tg}\frac{\gamma}{2};$$

$$S = p(p-a)\operatorname{tg}\frac{\alpha}{2} = p(p-b)\operatorname{tg}\frac{\beta}{2} = p(p-c)\operatorname{tg}\frac{\gamma}{2};$$

$$S = \frac{abc}{4R}; \quad S = pr; \quad S = \sqrt{rr_a r_b r_c}.$$

4.2. Вычисление элементов треугольника. Треугольник можно задать либо тремя сторонами, либо одной стороной и двумя углами, либо двумя сторонами и углом между ними. В качестве тройки элементов, задающих треугольник, можно выбрать и некоторые другие наборы элементов треугольника. Например, треугольник можно задать основанием a, высотой h_a и углом при основании. Имея тройку элементов, задающих треугольник, с помощью теоремы синусов и теоремы косинусов можно вычислить все остальные элементы треугольника.

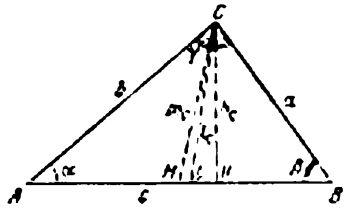


Рис. 6.23

Пример 1. Пусть треугольник ABC задан тремя сторонами a, b, c (рис. 6.23). Требуется вычислить все остальные элементы треугольника.

Углы α и β треугольника могут быть найдены с помощью теоремы косинусов:

$$a^2 = b^2 + c^2 - 2bc\cos\alpha \Rightarrow \cos\alpha = \frac{b^2+c^2-a^2}{2bc} \Rightarrow a = \arccos\frac{b^2+c^2-a^2}{2bc},$$

$$b^2 = a^2 + c^2 - 2ac\cos\beta \Rightarrow \cos\beta = \frac{a^2+c^2-b^2}{2ac} \Rightarrow \beta = \arccos\frac{a^2+c^2-b^2}{2ac}.$$

Третий угол γ можно вычислить, используя равенство, связывающее углы треугольника:

$$\alpha+\beta+\gamma = 180^\circ \Rightarrow \gamma = 180^\circ - \arccos\frac{b^2+c^2-a^2}{2bc} - \arccos\frac{a^2+c^2-b^2}{2ac}.$$

Высота треугольника h_c, опущенная на сторону c (см. рис. 6.23), может быть найдена из прямоугольного треугольника CHB:

$$\frac{h_c}{a} = \sin\beta \Rightarrow h_c = a\sin\beta = a\sin\left[\arccos\frac{a^2+c^2-b^2}{2ac}\right] \Rightarrow h_c = a\sqrt{1-\left(\frac{a^2+c^2-b^2}{2ac}\right)^2}.$$

Для нахождения медианы m_c воспользуемся тем, что точка M делит сторону c пополам ($|AM| = |MB|$). Следовательно, в треугольнике MBC оказываются известными две стороны $|MB| = c/2$, $|BC| = a$ и угол β между ними. Используя теорему косинусов для треугольника MBC, получаем

$$m_c^2 = \left(\frac{c}{2}\right)^2 + a^2 - ac\cos\beta = \left(\frac{c}{2}\right)^2 + a^2 - \frac{a^2+c^2-b^2}{2} = \frac{a^2+b^2}{2} - \frac{c^2}{4}.$$

Биссектрису l_c можно найти из треугольника LBC. В треугольнике LBC нам известны сторона BC ($|BC| = a$) и угол β. Угол $LCB = \gamma/2$ равен

$$\frac{\gamma}{2} = 90^\circ - \frac{1}{2}\arccos\frac{b^2+c^2-a^2}{2bc} - \frac{1}{2}\arccos\frac{a^2+c^2-b^2}{2ac}.$$

Отсюда

$$\angle CLB = 180^\circ - \angle LCB - \angle LBC = 180^\circ - \left[90^\circ - \frac{1}{2}\arccos\frac{b^2+c^2-a^2}{2bc} - \frac{1}{2}\arccos\frac{a^2+c^2-b^2}{2ac}\right] - \arccos\frac{a^2+c^2-b^2}{2ac} = 90^\circ + \frac{1}{2}\arccos\frac{b^2+c^2-a^2}{2bc} - \frac{1}{2}\arccos\frac{a^2+c^2-b^2}{2ac}.$$

Воспользуемся теперь теоремой синусов для нахождения биссектрисы l_c:

$$\frac{l_c}{\sin\beta}=\frac{b}{\sin\angle CLB}\Rightarrow l_c=\frac{a\sin\beta}{\sin\angle CLB}.$$

Вычисление синуса угла CLB будет довольно трудоемкой задачей.

Биссектрису треугольника можно вычислить и другим способом. Как известно, биссектриса делит сторону треугольника в отношении, пропорциональном прилежащим сторонам:

$$\frac{|AL|}{|AC|}=\frac{|LB|}{|CB|}.$$

Обозначив $|LB|=x$, запишем это равенство в виде

$$\frac{c-x}{b}=\frac{x}{a}\Rightarrow ac-ax=bx\Rightarrow ax=x(a+b)\Rightarrow x=\frac{ac}{a+b}.$$

Теперь в треугольнике LBC нам известны две стороны $|BC|=a$, $|LB|=\frac{ac}{a+b}$ и угол β между ними. Применяя теорему косинусов, получаем

$$l_c^2=|BC|^2+|LB|^2-2|BC|\cdot|LB|\cos\beta\Rightarrow l_c^2=$$
$$=a^2+\left(\frac{ac}{a+b}\right)^2-\frac{2a^2c}{a+b}\cos\beta\Rightarrow l_c^2=$$
$$=a^2+\left(\frac{ac}{a+b}\right)^2-\frac{2a^2c}{a+b}\cdot\frac{a^2+c^2-b^2}{2ac}\Rightarrow l_c=$$
$$=\sqrt{a^2+\left(\frac{ac}{a+b}\right)^2-\frac{a(a^2+c^2-b^2)}{a+b}}.$$

Площадь треугольника ABC может быть найдена, например, по формуле Герона:

$$S_{\triangle ABC}=\sqrt{p(p-a)(p-b)(p-c)},$$

а радиусы вписанной и описанной окружностей — с помощью формул

$$S_{\triangle ABC}=pr\Rightarrow r=\frac{S_{\triangle ABC}}{p},\quad S_{\triangle ABC}=\frac{abc}{4R}\Rightarrow R=\frac{abc}{4S_{\triangle ABC}}.$$

Пример 2. Рассмотрим теперь треугольник, у которого даны сторона a и два угла α и β (рис. 6.24).

Третий угол γ можно найти из равенства, связывающего углы треугольника: $\alpha+\beta+\gamma=180^\circ\Rightarrow\gamma=180^\circ-\alpha-\beta$. Стороны AB и AC находятся с помощью теоремы синусов:

$$\frac{|AB|}{\sin\gamma}=\frac{a}{\sin\alpha}\Rightarrow\frac{|AB|}{\sin(180^\circ-\alpha-\beta)}=\frac{a}{\sin\alpha}\Rightarrow$$
$$\Rightarrow\frac{|AB|}{\sin(\alpha+\beta)}=\frac{a}{\sin\alpha}\Rightarrow|AB|=\frac{a\sin(\alpha+\beta)}{\sin\alpha},$$
$$\frac{|AC|}{\sin\beta}=\frac{a}{\sin\alpha}\Rightarrow|AC|=\frac{a\sin\beta}{\sin\alpha}.$$

Таким образом, найдены три стороны; все остальные элементы треугольника (медиану, высоту, биссектрису и т. д.) можно найти

так же, как и в предыдущей задаче. Заметим, что радиус описанной окружности при данном наборе элементов, определяющих треугольник, удобно находить по формуле

$$\frac{a}{\sin\alpha}=2R \Rightarrow R=\frac{a}{2\sin\alpha}.$$

Пример 3. Если в треугольнике даны две стороны b, c и угол α между ними (рис. 6.25), то с помощью теоремы косинусов

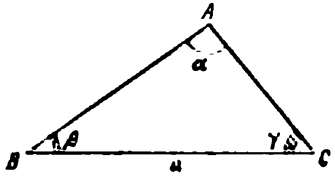


Рис. 6.24

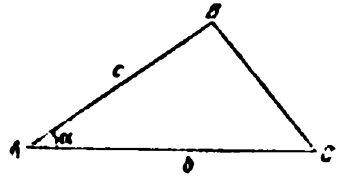


Рис. 6.25

сразу можно найти третью сторону:

$$|BC|^2=b^2+c^2-2bc\cos\alpha,$$

после чего все остальные элементы треугольника (углы, высоты, медианы, биссектрисы и т. д.) находятся так же, как и в примере 1.

Глава 7

МЕТОД КООРДИНАТ

Понятие прямоугольной системы координат на плоскости впервые появилось в геометрии еще до начала нашей эры. С ее помощью математик Александрийской школы Аполлоний определял и изучал кривые второго порядка — эллипс, гиперболу и параболу. В XVII веке в работах французских математиков Р. Декарта и П. Ферма были введены правила выбора знаков в прямоугольной системе координат и заложены основы аналитической геометрии на плоскости — раздела математики, устанавливающего связь между алгеброй и геометрией. Аналитическая геометрия сыграла важную роль в развитии понятия числа: благодаря правилу выбора знаков координат отрицательные числа, которые не признавало большинство математиков средневековья, получили наглядное изображение и окончательно утвердились в математике. В последующем применение прямоугольной декартовой системы координат сыграло решающую роль при утверждении в математике комплексных чисел.

§ 1. Системы координат

1.1. Координатная ось. *Координатной осью* называется прямая, на которой фиксированы две различные точки: точка O, называемая *началом координат*, и точка E, называемая *единичной точкой* (рис. 7.1). Обычно точку E располагают справа от точки O. Положительным направлением оси координат считается направление луча, выходящего

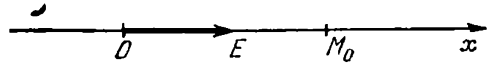


Рис. 7.1

из точки O и содержащего точку E. Противоположное направление считается отрицательным направлением оси координат. Отрезок OE называется *масштабным* или *единичным отрезком*. Координатную ось обычно обозначают Ox.

Вектор $\overrightarrow{OE}$ называют *единичным вектором* (или *ортом*) и обычно обозначают буквой $\boldsymbol{e}$.

Координатой точки M_0, лежащей на координатной оси, называется число x_0, определяемое равенством $x_0 = \pm \dfrac{|OM_0|}{|OE|}$, причем перед дробью берется знак плюс, когда точки E и M_0 лежат по одну сторону от точки O, и знак минус, когда точки E и M_0 расположены по разные стороны относительно точки O; если точка M_0 совпадает с точкой O, то $x_0 = 0$.

Расстояние d между точками M_1 и M_2 с координатами x_1 и x_2 вычисляется по формуле

$$d = |x_2 - x_1|.$$

1.2. Прямоугольная декартова система координат на плоскости. *Прямоугольной декартовой системой координат* на плоскости называется упорядоченная пара двух взаимно перпендикулярных координатных осей Ox и Oy, причем началом координат каждой из осей служит их общая точка O (начало прямоугольной декартовой системы координат) (рис. 7.2). Оси Ox и Oy упорядочены следующим образом: если ось Ox повернуть вокруг точки O на угол $\pi/2$ против движения часовой стрелки, то она совпадает с осью Oy. При этом масштабные отрезки OE_1 и OE_2 координатных осей Ox и Oy обычно выбираются так, чтобы их длины были равны: $|OE_1| = |OE_2|$. Тогда при повороте оси Ox на угол $\pi/2$ против движения часовой стрелки вокруг точки O точки E_1 и E_2 совместятся. Ось Ox называется *осью абсцисс*, ось Oy — *осью ординат*. Векторы $\boldsymbol{e}_1 = \overrightarrow{OE_1}$ и $\boldsymbol{e}_2 = \overrightarrow{OE_2}$ называются *базисными векторами* прямоугольной декартовой системы координат (или *ортами*) и обычно обозначаются буквами $\boldsymbol{i}$ и $\boldsymbol{j}$: $\boldsymbol{e}_1 = \boldsymbol{i}$, $\boldsymbol{e}_2 = \boldsymbol{j}$.

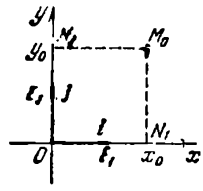


Рис. 7.2

Таким образом, можно считать, что прямоугольная декартова система координат на плоскости задается некоторой точкой O (началом координат) и упорядоченной парой взаимно перпендикулярных единичных векторов $(\boldsymbol{i}, \boldsymbol{j})$.

Координатные оси Ox и Oy разбивают плоскость на четыре четверти (*квадранта*). Часть плоскости, лежащая выше оси Ox и правее оси Oy, считается первым квадрантом; часть плоскости, лежащая выше оси Ox и левее оси Oy, — вторым; часть плоскости, лежащая ниже оси Ox и левее оси Oy, — третьим, и часть плоскости, лежащая ниже оси Ox и правее оси Oy, — четвертым квадрантом. Плоскость с построенной системой координат называется *координатной плоскостью*.

При изложенном выше способе упорядоченности координатных осей систему координат называют *правой* прямоугольной декартовой системой координат в отличие от *левой* прямоугольной декартовой системы координат, в которой оси упорядочены следующим образом: первая ось (ось Ox) совмещается со второй осью (осью Oy) поворотом на угол $\pi/2$ по движению часовой стрелки. Далее под словами «прямоугольная декартова система координат» будет пониматься правая прямоугольная декартова система координат.

В физике и механике наряду с декартовыми системами координат употребляются *косоугольные* (правые и левые) системы координат, в которых совмещение осей происходит при повороте на угол, отличный от прямого, а также *криволинейные* системы координат (полярные, сферические, цилиндрические и т. д.).

Координаты точки на плоскости. Пусть Oxy — прямоугольная декартова система координат на плоскости с началом O (см. рис. 7.2), M_0 — некоторая точка плоскости. Опустим из точки M_0 на оси Ox и Oy перпендикуляры, которые пересекут оси в точках N_1 и N_2 соответственно. Обозначим координату точки N_1 (лежащей на оси Ox) через x_0, а координату точки N_2 (лежащей на оси Oy) — через y_0.

Координатами точки M_0 в прямоугольной декартовой системе координат называется упорядоченная пара чисел $(x_0; y_0)$. Число x_0 называют *абсциссой*, число y_0 — *ординатой* точки M_0 и пишут $M_0(x_0; y_0)$.

Расстояние между точками A и B, имеющими координаты $(x_1; y_1)$ и $(x_2; y_2)$ соответственно, вычисляется по формуле

$$|AB| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Преобразования прямоугольных координат при параллельном переносе осей. Пусть Oxy и $O'x'y'$ — две прямоугольные декартовы системы координат на плоскости с началом координат O и O' $(O \neq O')$ соответственно, с одинаково

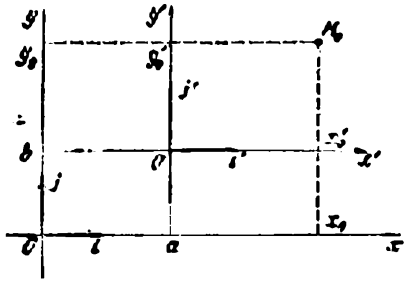


Рис. 7.3

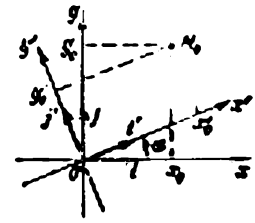


Рис. 7.4

направленными осями (рис. 7.3) и одинаковыми масштабными отрезками. Пусть точка O' относительно системы координат Oxy имеет координаты $(a; b)$. Число a называют величиной *сдвига* системы координат $O'x'y'$ относительно системы Oxy по направлению оси Ox, а число b — величиной *сдвига* по направлению оси Oy.

Пусть точка M_0 имеет координаты $(x_0; y_0)$ относительно системы Oxy. Тогда ее координаты $(x_0'; y_0')$ относительно системы $O'x'y'$ связаны с координатами относительно системы Oxy формулами

$$x_0' = x_0 - a, \quad y_0' = y_0 - b.$$

Эти формулы называют *формулами преобразования* прямоугольных координат точки M_0 при параллельном переносе осей.

Правило преобразования координат точки M_0 при переходе от одной прямоугольной системы координат к другой можно сформулировать следующим образом.

При параллельном переносе прямоугольной системы координат на величину a в направлении оси Ox и на величину b в направлении оси Oy абсциссы всех точек уменьшаются или увеличиваются (в зависимости от знака величины a) на величину a, ординаты — на величину b.

Формулы преобразования прямоугольных координат при повороте осей. Пусть Oxy и $Ox'y'$ — две прямоугольные декартовы системы координат, причем система координат $Ox'y'$ получена из системы Oxy *поворотом* на угол α около общего начала координат O (рис. 7.4). Пусть точка M_0 имеет координаты $(x_0; y_0)$ относительно системы координат Oxy и координаты $(x_0'; y_0')$ относительно системы координат $Ox'y'$. Координаты точки M_0

в системах координат Oxy и $Ox'y'$ связаны формулами

$$\begin{aligned} x_0 &= x_0' \cos\alpha - y_0' \sin\alpha, \\ y_0 &= x_0' \sin\alpha + y_0' \cos\alpha \end{aligned} \tag{1}$$

или равносильными формулами

$$\begin{aligned} x_0' &= x_0 \cos\alpha + y_0 \sin\alpha, \\ y_0' &= -x_0 \sin\alpha + y_0 \cos\alpha. \end{aligned}$$

Формулы преобразования прямоугольных координат при параллельном переносе и повороте осей. Пусть Oxy и $O'x'y'$ — две прямоугольные декартовы системы координат, и пусть система координат $O'x'y'$ получена из системы Oxy параллельным переносом с величиной сдвига соответственно a и b и с поворотом вокруг точки O' на угол α (рис. 7.5).

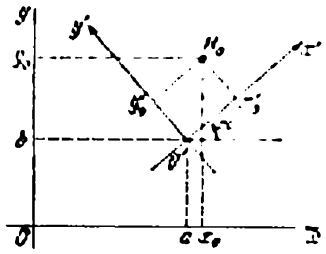


Рис. 7.5

Пусть M_0—точка плоскости с координатами $(x_0; y_0)$ относительно системы координат Oxy и координатами $(x_0'; y_0')$ относительно системы координат $O'x'y'$. Координаты точки M_0 относительно систем координат Oxy и $O'x'y'$ связаны преобразования формулами

$$\begin{aligned} x_0 &= x_0' \cos\alpha - y_0' \sin\alpha + a, \\ y_0 &= x_0' \sin\alpha + y_0' \cos\alpha + b, \end{aligned} \tag{2}$$

выражающими координаты точки M_0 относительно системы координат Oxy через координаты той же самой точки M_0 относительно системы $O'x'y'$.

Так как M_0 — произвольная точка, то индекс 0 у переменных x и y обычно опускают и записывают формулы (2) в виде

$$\begin{aligned} x &= x' \cos\alpha - y' \sin\alpha + a, \\ y &= x' \sin\alpha + y' \cos\alpha + b. \end{aligned} \tag{3}$$

Разрешая соотношения (3) относительно величин x' и y', получаем выражения для координат точки относительно системы $O'x'y'$ через координаты той же точки относительно системы Oxy:

$$\begin{aligned} x' &= (x-a)\cos\alpha + (y-b)\sin\alpha, \\ y' &= -(x-a)\sin\alpha + (y-b)\cos\alpha. \end{aligned}$$

1.3. Полярная система координат. Связь между прямоугольными и полярными координатами. *Полярная система координат* на плоскости задается выбором точки O, называемой *полюсом*, и направленным лучом Ox (*полярной осью*) с выбранной на нем единичной точкой E. Отрезок OE называется *единичным* или масштабным отрезком (рис. 7.6).

Точке M_0, отличной от точки O, ставится в соответствие упорядоченная пара чисел $(\rho_0; \varphi_0)$, называемая *полярными координатами* точки M_0. Число ρ_0, называемое *полярным радиусом* точки M_0, представляет собой длину отрезка OM_0. Число φ_0, называемое *полярным углом*, есть радианная величина угла EOM_0 (см. рис. 7.6).

Между множеством всех точек плоскости (за исключением точки O) и множеством упорядоченных пар чисел $(\rho_0; \varphi_0)$, где $\varphi_0 \in [0; 2\pi)$,

существует взаимно однозначное соответствие. Для точки O (полюса) величина полярного угла не определена.

Связь между полярными и прямоугольными координатами. Если за полюс принять начало прямоугольной декартовой системы координат, а за полярную ось — положительное направление оси Ox (рис. 7.7), то полярные координаты $(\rho_0; \varphi_0)$

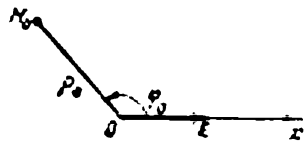


Рис. 7.6

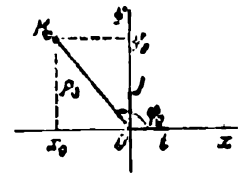


Рис. 7.7

точки M_0, не совпадающей с полюсом, и ее прямоугольные координаты $(x_0; y_0)$ связаны равенствами

$$x_0 = \rho_0 \cos \varphi_0, \quad y_0 = \rho_0 \sin \varphi_0.$$

Эти формулы называют *формулами перехода* от полярных координат точки M_0 к прямоугольным координатам точки M_0.

Формулы перехода от прямоугольных координат точки M_0 $(x_0; y_0)$ к ее полярным координатам: полярный радиус ρ_0 вычисляется по формуле

$$\rho_0 = \sqrt{x_0^2 + y_0^2},$$

а полярный угол φ_0 вычисляется из равенств

$$\sin \varphi_0 = y_0 / \sqrt{x_0^2 + y_0^2}, \quad \cos \varphi_0 = x_0 / \sqrt{x_0^2 + y_0^2}$$

при условии, что $\varphi_0 \in [0; 2\pi)$.

1.4. Прямоугольная декартова система координат в пространстве. Пусть α — некоторая плоскость в пространстве с выбранной на ней

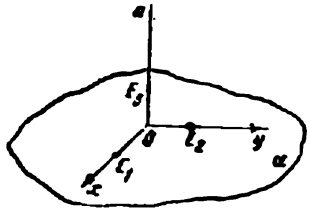


Рис. 7.8

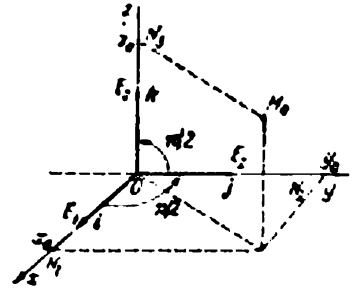


Рис. 7.9

прямоугольной декартовой системой координат Oxy. Проведем через точку O прямую a, перпендикулярную плоскости α (рис. 7.8). Выберем на прямой a направление и масштабный отрезок OE_3, равный масштабным отрезкам OE_1 и OE_2 координатных осей Ox и Oy соответственно (на рис. 7.8 положительное направление на прямой a отмечено стрелкой). Прямая a с выбранным направлением и масштабным отрезком — третья координатная ось, которую обозначим Oz.

Прямоугольной декартовой системой координат в пространстве называется упорядоченная тройка взаимно перпендикулярных координатных осей Ox, Oy и Oz. Прямоугольная декартова система координат $Oxyz$ называется *правой* (рис. 7.9), если Oxy — правая прямоугольная система координат, а ось Oy совмещается с осью Oz поворотом на угол $\pi/2$ против движения часовой стрелки. Прямоугольная декартова система координат $Oxyz$ называется *левой*, если система координат Oxy — левая, либо, если Oxy — правая, ось Oy совмещается с осью Oz поворотом на угол $\pi/2$ по движению часовой стрелки.

Ось Ox называется *осью абсцисс*, ось Oy — *осью ординат*, ось Oz — *осью аппликат*. Плоскости, проходящие через каждые две координатные оси, называют *координатными плоскостями*; пространство, в котором выбрана система координат, называют *координатным пространством*.

Векторы $\overrightarrow{OE_1}$, $\overrightarrow{OE_2}$, $\overrightarrow{OE_3}$ называются *базисными векторами* прямоугольной декартовой системы координат (или ортами) и обычно обозначаются соответственно $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$.

Координаты точки в пространстве. Пусть M_0 — некоторая точка пространства, а $Oxyz$ — некоторая пространственная прямоугольная декартова система координат с началом в точке O (см. рис. 7.9). Построим плоскость, содержащую точку M_0 и параллельную координатной плоскости Oyz. Она пересечет координатную ось Ox в некоторой точке N_1. Координату точки N_1 обозначим через x_0. Построив плоскости, параллельные координатным плоскостям Oxz и Oxy и содержащие точку M_0, получаем точки N_2 и N_3 — точки пересечения построенных плоскостей с осями Oy и Oz соответственно. Координаты точек N_2 и N_3, принадлежащих соответственно осям Oy и Oz, обозначим через y_0 и z_0.

Координатами точки M_0 относительно прямоугольной декартовой системы координат $Oxyz$ называют упорядоченную тройку чисел $(x_0; y_0; z_0)$ и пишут $M_0(x_0; y_0; z_0)$. Число x_0 называется *абсциссой* точки M_0, y_0 — ее *ординатой*, z_0 — *аппликатой*.

Расстояние между точками $A(x_1; y_1; z_1)$ и $B(x_2; y_2; z_2)$ вычисляется по формуле

$$|AB| = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}.$$

§ 2. Линии на плоскости и поверхности в пространстве

2.1. Линии первого и второго порядка. Пусть $P(x; y)$ — многочлен двух переменных x и y. Уравнение

$$P(x, y)=0 \qquad (1)$$

в прямоугольной декартовой системе координат Oxy задает множество точек, которое будем называть *линией*.

Если $P(x, y)$ — многочлен первой степени, то линия, определяемая уравнением (1), называется *линией первого порядка*, а уравнение (1) может быть записано в виде

$$Ax+By+C=0.$$

Линия первого порядка — это прямая.

Если $P(x, y)$ — многочлен второй степени, то линия, определяемая уравнением (1), называется *линией второго порядка*, а уравнение (1)

может быть записано в виде

$$Ax^2+Bxy+Cy^2+Dx+Ky+L=0. \qquad (2)$$

Линия второго порядка, задаваемая уравнением (2), в зависимости от значений коэффициентов A, B, C, D, K, L определяет эллипс, гиперболу или параболу, а при некоторых значениях коэффициентов — точку или две прямые (последние случаи называют *вырожденными*).

Преобразованием координат

$$x=x'\cos\alpha-y'\sin\alpha+x_0,$$
$$y=x'\sin\alpha+y'\cos\alpha+y_0$$

при соответствующем подборе значений α, x_0, y_0 уравнение (2) в зависимости от значений коэффициентов приводится к одному из следующих уравнений (исключая вырожденные случаи):

$$\frac{x'^2}{a^2}+\frac{y'^2}{b^2}=1 \qquad (эллипс),$$

$$\frac{x'^2}{a^2}-\frac{y'^2}{b^2}=1 \qquad (гипербола),$$

$$y'^2=2px' \qquad (парабола).$$

Более подробно о линиях второго порядка см. п. 2.4—2.6.

2.2. Прямая. Каждое линейное уравнение с двумя переменными x и y, т. е. уравнение вида

$$Ax+By+C=0, \qquad (3)$$

где A и B не равны нулю одновременно, в прямоугольной декартовой системе координат Oxy задает прямую. Обратно, каждая прямая в прямоугольной декартовой системе координат может быть задана линейным уравнением вида (3).

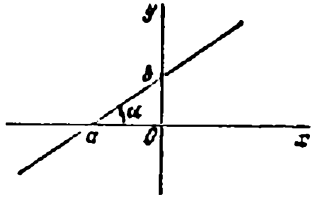


Рис. 7.10

Уравнение вида (3) называется *общим уравнением прямой*.

Некоторые частные формы записи уравнения прямой.

1. Прямая, образующая угол α с положительным направлением оси Ox и пересекающая ось Oy в точке $(0; b)$ (рис. 7.10), задается уравнением

$$y=kx+b, \text{ где } \quad k=\operatorname{tg}\alpha, \quad \alpha\in[0;\ \pi/2)\cup(\pi/2;\ \pi).$$

Коэффициент k называется *угловым коэффициентом* прямой.

Уравнение прямой, образующей угол α с положительным направлением оси Ox и проходящей через точку $(x_0; y_0)$.

$$y-y_0=k(x-x_0), \quad k=\operatorname{tg}\alpha.$$

2. Уравнение прямой в отрезках. Уравнение прямой, пересекающей ось Ox в точке $(a; 0)$ и ось Oy в точке $(0; b)$ (см. рис. 7.10):

$$\frac{x}{a}+\frac{y}{b}=1, \quad a\neq 0, \quad b\neq 0.$$

3. Уравнение прямой, проходящей через две (не совпадающие) точки $M_1(x_1; y_1)$ и $M_2(x_2; y_2)$:

$$\frac{y-y_1}{y_2-y_1}=\frac{x-x_1}{x_2-x_1}.$$

Взаимное расположение трех точек. Необходимым и достаточным условием принадлежности трех точек $M_1(x_1; y_1)$, $M_2(x_2; y_2)$, $M_3(x_3; y_3)$ одной прямой является равенство нулю определителя третьего порядка:

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}=0.$$

Координаты точки M_0, делящей отрезок M_1M_2 с концами $M_1(x_1; y_1)$, $M_2(x_2; y_2)$ в отношении λ, определяются соотношениями

$$x_0=\frac{x_1+\lambda x_2}{1+\lambda}, \quad y_0=\frac{y_1+\lambda y_2}{1+\lambda}.$$

Расстояние от точки до прямой. Расстояние d от точки $M_0(x_0; y_0)$ до прямой, заданной в прямоугольной системе координат уравнением $Ax+By+C=0$, вычисляется по формуле

$$d=\frac{|Ax_0+By_0+C|}{\sqrt{A^2+B^2}}.$$

Взаимное расположение двух прямых на плоскости. Две прямые, заданные уравнениями

$$A_1x+B_1y+C_1=0, \quad A_2x+B_2y+C_2=0,$$

пересекаются тогда и только тогда, когда отличен от нуля определитель второго порядка, составленный из коэффициентов при переменных x и y:

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}\neq 0.$$

Координаты $(x_0; y_0)$ точки пересечения двух данных прямых определяются по формулам, называемым *формулами Крамера*:

$$x_0=\frac{\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}, \quad y_0=-\frac{\begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}.$$

Для того чтобы две прямые, заданные уравнениями

$$A_1x+B_1y+C_1=0, \quad A_2x+B_2y+C_2=0,$$

были *взаимно перпендикулярны*, необходимо и достаточно, чтобы коэффициенты A_1, B_1, A_2, B_2 удовлетворяли уравнению

$$A_1A_2+B_1B_2=0.$$

Для того чтобы две прямые, заданные уравнениями с угловыми коэффициентами k_1 и k_2:

$$y=k_1x+b_1, \quad y=k_2x+b_2,$$

были *взаимно перпендикулярны*, необходимо и достаточно, чтобы k_1 и k_2 были связаны между собой равенством

$$k_1k_2=-1,$$

Угол между прямыми, задаваемыми уравнениями

$$y=k_1x+b_1, \quad y=k_2x+b_2,$$

определяется равенством

$$\operatorname{tg}\alpha=\left|\frac{k_1-k_2}{1+k_1k_2}\right|,$$

где $k_1k_2\neq-1$, т. е. прямые не являются взаимно перпендикулярными. При $k_1=k_2$ $\alpha=0$ и прямые *параллельны*.

2.3. **Окружность.** *Уравнение окружности* радиуса R с центром в начале O прямоугольной декартовой системы координат Oxy:

$$x^2+y^2=R^2.$$

Уравнение окружности радиуса R с центром в точке с координатами

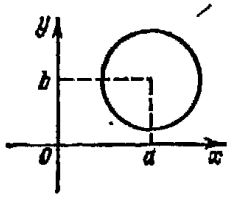
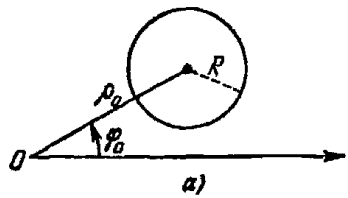


Рис. 7.11



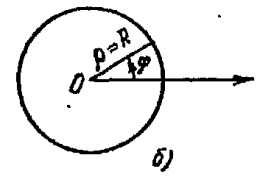


Рис. 7.12

(a; b) в прямоугольной декартовой системе координат (рис. 7.11):

$$(x-a)^2+(y-b)^2=R^2.$$

Уравнение окружности радиуса R с центром в точке $(\rho_0; \varphi_0)$ в полярной системе координат (рис. 7.12, *а*):

$$\rho^2-2\rho\rho_0\cos(\varphi-\varphi_0)+\rho_0^2=R^2.$$

Если центр окружности совпадает с полюсом (рис. 7.12, *б*), то уравнение окружности радиуса R в полярной системе координат принимает вид

$$\rho=R.$$

2.4. **Эллипс.** Множество точек плоскости, для которых сумма расстояний от двух данных точек плоскости F_1 и F_2, называемых *фокусами*, есть положительная постоянная величина, называется *эллипсом*. Для того чтобы указанное множество точек плоскости не было пустым множеством, необходимо, чтобы эта постоянная величина была больше расстояния между фокусами.

Каноническое уравнение эллипса. Пусть на плоскости выбрана прямоугольная декартова система координат Oxy (рис. 7.13), и пусть точки F_1 и F_2 с координатами $(-c; 0)$ и $(c; 0)$ соответственно — фокусы эллипса. Если M — произвольная точка эллипса, координаты которой будем обозначать $(x; y)$, то, по определению эллипса, имеет место равенство *)

$$|F_1M|+|F_2M|=2a. \tag{4}$$

*) Здесь удобно указанную постоянную величину обозначать через $2a$.

Отрезки F_1M и F_2M называются *фокальными радиусами* эллипса; их длины вычисляются по формулам

$$|F_1M| = \sqrt{(x+c)^2+y^2}, \quad |F_2M| = \sqrt{(x-c)^2+y^2}.$$

Подставляя выражения для длин фокальных радиусов в равенство (4), после несложных преобразований получаем следующую запись уравнения (4):

$$(a^2-c^2)\,x^2+a^2y^2=a^2(a^2-c^2). \qquad (5)$$

Вводя новую величину $b=\sqrt{a^2-c^2}$, приводим уравнение (5) к виду

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1. \qquad (6)$$

Уравнение (6) называется *каноническим уравнением* эллипса.

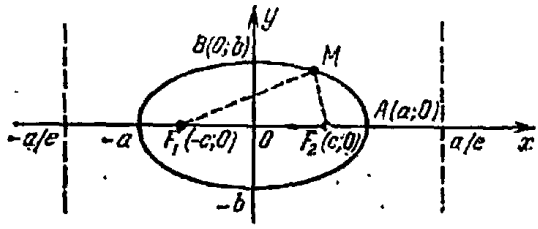


Рис. 7.13

Точки пересечения эллипса с его осями симметрии (в нашем случае осями симметрии являются координатные оси Ox и Oy) называются *вершинами* эллипса. На рис. 7.13 вершины эллипса — точки с координатами $(a; 0)$, $(-a; 0)$, $(0; b)$, $(0; -b)$. Отрезок OA ($|OA|=a$) называется *большой полуосью* эллипса, а отрезок OB ($|OB|=b$) — *малой полуосью* эллипса. Точка O (в нашем случае начало системы координат) называется *центром* эллипса. Эллипс является центрально симметричной фигурой относительно точки O.

Эксцентриситетом эллипса называется отношение расстояния между фокусами эллипса к длине его большой оси:

$$e=\frac{c}{a}<1 \quad \text{или} \quad e=\sqrt{1-\left(\frac{b}{a}\right)^2}. \qquad (7)$$

Эксцентриситет характеризует форму эллипса: чем ближе эксцентриситет к единице, тем меньше отношение b/a и тем больше эллипс вытянут; чем меньше эксцентриситет, тем ближе отношение b/a к единице и тем больше эллипс «похож» на окружность.

Используя определение эксцентриситета эллипса (7), уравнение (6) и равенство $b^2=a^2-c^2$, можно получить выражения фокальных радиусов эллипса через его эксцентриситет:

$$|F_1M|=a+ex; \quad |F_2M|=a-ex.$$

Директрисы эллипса. Предположим, что эллипс, задаваемый каноническим уравнением (6), вытянут в направлении оси Ox, т. е. $a>b$. Две прямые, перпендикулярные к большой оси эллипса (в нашем случае к оси Ox) и расположенные на расстоянии a/e от центра эллипса (в нашем случае от точки O), называются *директри-*

сами эллипса (см. рис. 7.13). Уравнения директрис эллипса имеют вид

$$x = -\frac{a}{e}, \qquad x = \frac{a}{e}.$$

Так как для эллипса $e < 1$, то правая директриса расположена правее правой вершины эллипса; аналогично левая директриса расположена левее его левой вершины.

Свойство директрис эллипса: если r — расстояние от произвольной точки эллипса до какого-нибудь фокуса, d — расстояние от той же точки до соответствующей этому фокусу директрисы, то отношение r/d есть постоянная величина, равная эксцентриситету эллипса:

$$\frac{r}{d} = e.$$

Уравнение касательной к эллипсу, проходящей через точку M_0 $(x_0; y_0)$, принадлежащую эллипсу, имеет вид

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} = 1.$$

2.5. Гипербола. Множество точек, для которых абсолютная величина разности расстояний от двух данных точек плоскости F_1 и F_2,

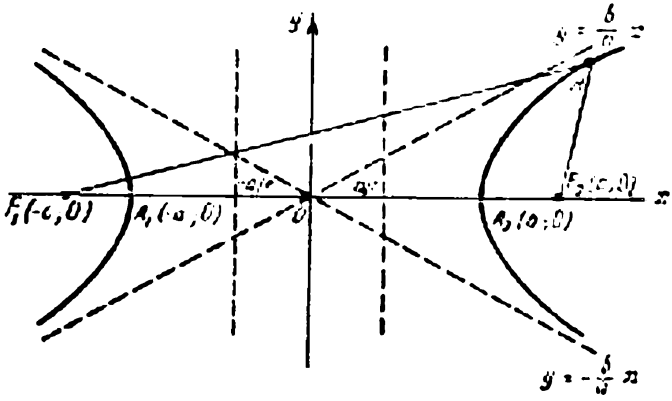


Рис. 7.11

называемых *фокусами*, есть постоянная положительная величина, называется *гиперболой*.

Множество точек M, образующих гиперболу, в соответствии с определением, удовлетворяет уравнению *)

$$\big|\,|F_1M| - |F_2M|\,\big| = 2a.$$

Отрезки F_1M и F_2M называются *фокальными радиусами* гиперболы.

Каноническое уравнение гиперболы. Пусть на плоскости выбрана прямоугольная декартова система координат Oxy (рис. 7.14), и пусть точки с координатами $(-c; 0)$ и $(c; 0)$ — фокусы

*) Постоянную величину удобно обозначить через $2a$.

гиперболы. Координаты точки $M(x; y)$, принадлежащей гиперболе, удовлетворяют уравнению

$$\left|\sqrt{(x+c)^2+y^2}-\sqrt{(x-c)^2+y^2}\right|=2a. \tag{8}$$

Избавляясь от иррациональностей, уравнение (8) можно привести к виду

$$(c^2-a^2)\,x^2-a^2y^2=a^2\,(c^2-a^2).$$

Обозначив $\sqrt{c^2-a^2}=b$ $(c>a)$, приведем это уравнение к виду

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1. \tag{9}$$

Уравнение (9) называется *каноническим уравнением* гиперболы.

Точки пересечения гиперболы с ее осями симметрии (в нашем случае осями симметрии являются координатные оси Ox и Oy) называются *вершинами* гиперболы (на рис. 7.14 — точки A_1 и A_2 с координатами $(-a; 0)$ и $(a; 0)$ соответственно). Точка O (в нашем случае начало координат) называется *центром* гиперболы. Гипербола является как центрально симметричной фигурой, так и фигурой, симметричной относительно осей Ox и Oy. *Ветви* гиперболы, лежащие в первой и третьей координатных четвертях (см. рис. 7.14), при $|x|\longrightarrow\infty$ приближаются к прямой $y=\frac{b}{a}x$. Ветви гиперболы, лежащие во второй и четвертой координатных четвертях, при $|x|\longrightarrow\infty$ приближаются к прямой $y=-\frac{b}{a}x$.

Прямоугольник со сторонами $2a$ и $2b$, расположенный симметрично относительно осей гиперболы, называется *основным прямоугольником* гиперболы (рис. 7.15). В математической литературе принято также называть *осями* гиперболы отрезки длиной $2a$ и $2b$, соединяющие середины противоположных сторон основного прямоугольника. В соответствии с этой терминологией говорят, что уравнение

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$$

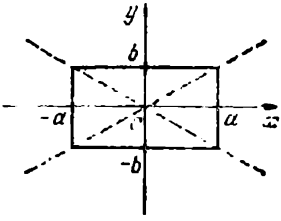


Рис. 7.15

определяет гиперболу с полуосями a и b.

Эксцентриситетом гиперболы называется отношение расстояния между фокусами этой гиперболы к расстоянию между ее вершинами:

$$e=\frac{c}{a}>1 \quad \text{или} \quad e=\sqrt{1+\left(\frac{b}{a}\right)^2}. \tag{10}$$

Эксцентриситет гиперболы характеризует форму ее основного прямоугольника, а значит, и форму самой гиперболы: чем меньше эксцентриситет гиперболы, тем больше вытянут ее основной прямоугольник в направлении оси, соединяющей вершины.

Гипербола называется *равносторонней*, если стороны основного прямоугольника равны.

Используя определение эксцентриситета гиперболы (10), уравнение гиперболы (9) и равенство $b^2=c^2-a^2$, можно получить

выражения фокальных радиусов через эксцентриситет:

$$|F_1M| = |a + ex|, \quad |F_2M| = |a - ex|.$$

Директрисы гиперболы. Две прямые, перпендикулярные к той оси гиперболы, которая ее пересекает, и расположенные на расстоянии a/e от центра гиперболы, называются *директрисами* гиперболы. Уравнения директрис в прямоугольной системе координат, относительно которой гипербола задается каноническим уравнением (9), имеют вид $x = -a/e$, $x = a/e$. Так как для гиперболы $e > 1$, то правая директриса расположена между центром и правой вершиной гиперболы; аналогично левая директриса расположена между центром и левой вершиной (см. рис. 7.14).

Свойство директрис гиперболы: если r — расстояние от произвольной точки гиперболы до какого-нибудь фокуса, d — расстояние от той же точки до соответствующей этому фокусу директрисы, то отношение r/d есть постоянная величина, равная эксцентриситету гиперболы: $r/d = e$.

Уравнение *касательной к гиперболе*, проходящей через точку $M_0(x_0; y_0)$, принадлежащую гиперболе, имеет вид

$$\frac{x_0x}{a^2} - \frac{y_0y}{b^2} = 1.$$

2.6. Парабола. Множество точек плоскости, равноотстоящих от заданной точки F, называемой *фокусом*, и от заданной прямой, называемой *директрисой*, называется *параболой* *).

Расстояние p от фокуса F до директрисы параболы называется *параметром* параболы.

Множество точек M, образующих параболу, в согласии с данным определением параболы, удовлетворяет уравнению

$$|FM| = d.$$

Каноническое уравнение параболы. Пусть на плоскости выбрана прямоугольная декартова система координат Oxy, и пусть точка с координатами $(p/2; 0)$ — фокус параболы, а прямая, задаваемая уравнением $x = -p/2$, — директриса параболы (рис. 7.16).

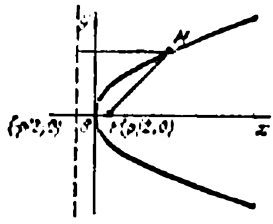


Рис. 7.16

Координаты $(x; y)$ любой точки M, принадлежащей параболе, удовлетворяют уравнению

$$\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = x + \frac{p}{2}. \qquad (11)$$

Избавляясь от иррациональности, приводим уравнение (11) к виду

$$y^2 = 2px. \qquad (12)$$

Уравнение (12) называется *каноническим уравнением параболы*.

Ось симметрии параболы (в нашем случае ось Ox) называется *осью* параболы. Точка пересечения параболы с осью симметрии называется *вершиной* параболы (в нашем случае вершина параболы совпадает с началом координат).

---

*) При этом предполагается, что точка F не принадлежит директрисе.

Заметим, что, в соответствии с определением параболы и свойствами директрис эллипса и гиперболы, можно условно считать, что парабола имеет эксцентриситет, равный единице.

Уравнение касательной к параболе, проходящей через точку $M_0(x_0; y_0)$, принадлежащую параболе, имеет вид

$$y_0 y = p(x + x_0).$$

2.7. Плоскость. Прямые в пространстве. В общем случае поверхность в пространстве в прямоугольной декартовой системе координат $Oxyz$ может быть задана уравнением

$$F(x, y, z) = 0. \tag{13}$$

В случае, когда $F(x, y, z)$ — многочлен первой степени, уравнение (13) определяет *плоскость*. Если $F(x, y, z)$ — многочлен второй степени (по совокупности переменных), то уравнение (13) определяет в зависимости от значений своих коэффициентов следующие поверхности второго порядка: эллипсоид (частным случаем которого является шар), двуполостный гиперболоид, однополостный гиперболоид, эллиптический гиперболоид и гиперболический параболоид, а также конус и цилиндр.

Плоскость. *Общее уравнение* плоскости в декартовой системе координат:

$$Ax + By + Cz + D = 0, \quad A^2 + B^2 + C^2 \neq 0.$$

Уравнение в отрезках:

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1, \quad abc \neq 0,$$

$(a; 0; 0)$, $(0; b; 0)$, $(0; 0; c)$ — точки пересечения плоскости с осями Ox, Oy и Oz соответственно.

Нормальное уравнение плоскости:

$$x \cos\alpha + y \cos\beta + z \cos\gamma - p = 0,$$

$\cos\alpha$, $\cos\beta$, $\cos\gamma$ — компоненты вектора единичной длины, перпендикулярного плоскости, p — расстояние от начала координат до плоскости.

Коэффициенты общего и нормального уравнений плоскости связаны соотношениями

$$\cos\alpha = \lambda A, \quad \cos\beta = \lambda B, \quad \cos\gamma = \lambda C, \quad p = -\lambda D,$$

$$|\lambda| = \frac{1}{\sqrt{A^2 + B^2 + C^2}}$$

(знак λ противоположен знаку D).

Уравнение плоскости, проходящей через три заданные точки $(x_i; y_i; z_i)$ $(i = 1, 2, 3)$, не лежащие на одной прямой:

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0.$$

Уравнение плоскости, проходящей через заданную точку (x_0, y_0, z_0) и перпендикулярной вектору $n = (A; B; C)$:

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

Необходимое и достаточное условие параллельности плоскостей $A_i x + B_i y + C_i z + D_i = 0$ $(i = 1, 2)$:

$$A_1 = \lambda A_2, \quad B_1 = \lambda B_2, \quad C_1 = \lambda C_2.$$

Необходимое и достаточное условие перпендикулярности плоскостей:

$$A_1 A_2 + B_1 B_2 + C_1 C_2 = 0.$$

Расстояние d от точки $(x_0; y_0; z_0)$ до плоскости $Ax + By + Cz + D = 0$:

$$d = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}}.$$

Расстояние d между двумя параллельными плоскостями $Ax + By + Cz + D_1 = 0$ и $Ax + By + Cz + D_2 = 0$:

$$d = \frac{|D_1 - D_2|}{\sqrt{A^2 + B^2 + C^2}}.$$

Угол φ *между плоскостями* $A_i x + B_i y + C_i z + D_i = 0$ $(i = 1, 2)$ определяется формулой

$$\cos\varphi = \frac{|A_1 A_2 + B_1 B_2 + C_1 C_2|}{\sqrt{A_1^2 + B_1^2 + C_1^2}\sqrt{A_2^2 + B_2^2 + C_2^2}}.$$

Прямые в пространстве. Прямая в пространстве может быть задана как линия пересечения двух непараллельных плоскостей

$$\begin{aligned} A_1 x + B_1 y + C_1 z + D_1 &= 0, \\ A_2 x + B_2 y + C_2 z + D_2 &= 0. \end{aligned} \tag{11}$$

Параметрическое уравнение прямой:

$$x = x_0 + r_x t, \quad y = y_0 + r_y t, \quad z = z_0 + r_z t,$$

$\boldsymbol{r} = (r_x, r_y, r_z)$ — *направляющий вектор прямой*, x_0, y_0, z_0 — координаты точки, принадлежащей прямой.

Компоненты направляющего вектора прямой связаны с коэффициентами системы (11) равенствами

$$\cos\alpha = \frac{r_x}{|\boldsymbol{r}|}, \quad r_x = (B_1 C_2 - B_2 C_1),$$

$$\cos\beta = \frac{r_y}{|\boldsymbol{r}|}, \quad r_y = (C_1 A_2 - C_2 A_1),$$

$$\cos\gamma = \frac{r_z}{|\boldsymbol{r}|}, \quad r_z = (A_1 B_2 - A_2 B_1).$$

Каноническое уравнение прямой, проходящей через точку (x_0, y_0, z_0) и имеющей направляющий вектор $\boldsymbol{r} = (r_x; r_y; r_z)$:

$$\frac{x - x_0}{r_x} = \frac{y - y_0}{r_y} = \frac{z - z_0}{r_z}.$$

Угол φ между двумя прямыми, заданными каноническими уравнениями

$$\frac{x - x_0}{r_x} = \frac{y - y_0}{r_y} = \frac{z - z_0}{r_z}, \qquad \frac{x - x_0'}{r_x'} = \frac{y - y_0'}{r_y'} = \frac{z - z_0'}{r_z'},$$

определяется формулой

$$\cos\varphi=\frac{|r_xr'_x+r_yr'_y+r_zr'_z|}{\sqrt{(r_x^2+r_y^2+r_z^2)(r'^2_x+r'^2_y+r'^2_z)}}.$$

Условие перпендикулярности двух прямых, заданных каноническими уравнениями:

$$r'_xr_x+r'_yr_y+r'_zr_z=0.$$

Условие параллельности двух прямых, заданных каноническими уравнениями:

$$\frac{r_x}{r'_x}=\frac{r_y}{r'_y}=\frac{r_z}{r'_z}.$$

Условие перпендикулярности прямой, заданной каноническими уравнениями, и плоскости $Ax+By+Cz+D=0$:

$$\frac{A}{r_x}=\frac{B}{r_y}=\frac{C}{r_z}.$$

Условие параллельности прямой, заданной каноническими уравнениями, и плоскости $Ax+By+Cz+D=0$:

$$Ar_x+Br_y+Cr_z=0.$$

Уравнение прямой, проходящей через точку $(x_0; y_0; z_0)$ и перпендикулярной плоскости $Ax+By+Cz+D=0$:

$$x-x_0=\lambda A,\quad y-y_0=\lambda B,\quad z-z_0=\lambda C.$$

Расстояние от точки $(x_1; y_1; z_1)$ до прямой, заданной каноническим уравнением:

$$d=\left[\frac{\begin{vmatrix} r_y & r_z\\ y_0-y_1 & z_0-z_1\end{vmatrix}^2+\begin{vmatrix} r_z & r_x\\ z_0-z & x_0-x_1\end{vmatrix}^2+\begin{vmatrix} r_x & r_y\\ x_0-x_1 & y_0-y_1\end{vmatrix}^2}{r_x^2+r_y^2+r_z^2}\right]^{1/2}.$$

§ 3. Векторы

3.1. Векторы. Основные понятия. Пусть A и B — две различные точки плоскости. Отрезок AB, у которого точку A считают началом, а точку B — концом, называют *вектором* AB и обозначают $\overrightarrow{AB}$. О векторе $\overrightarrow{AB}$ говорят также, что он приложен к точке A. Направление, определяемое лучом AB, называется *направлением* вектора $\overrightarrow{AB}$, а длина отрезка AB называется *длиной* (или *модулем*) вектора $\overrightarrow{AB}$. Длина (модуль) вектора $\overrightarrow{AB}$ обозначается $|\overrightarrow{AB}|$. На чертежах вектор $\overrightarrow{AB}$ обычно изображают прямолинейной стрелкой с началом в точке A и концом в точке B (рис. 7.17).

Рис. 7.17

Среди всех векторов с началом в точке A имеется один вектор, длина которого равна нулю. Считается, что этот вектор начинается и кончается в точке A. Вектор, начинающийся и кончающийся в

точке A, называется *нулевым* вектором и обозначается $\overrightarrow{AA}$. Понятие направления для нулевого вектора $\overrightarrow{AA}$ не вводится.

Определение пространственного вектора дословно совпадает с определением вектора на плоскости, при этом лишь считают, что точки A и B — две точки пространства.

Равенство векторов. Пусть $\overrightarrow{AB}$ и $\overrightarrow{CD}$ — два вектора плоскости (или пространства). Говорят, что вектор $\overrightarrow{AB}$ равен вектору $\overrightarrow{CD}$, и пишут $\overrightarrow{AB}=\overrightarrow{CD}$, если:

1) длина отрезка AB равна длине отрезка CD;

2) лучи AB и CD одинаково направлены.

При таком определении равенства векторов множество всех векторов, равных вектору $\overrightarrow{AB}$, называют *свободным* вектором. Свободные векторы обычно обозначают малыми латинскими буквами: $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{x}$, а их длины — соответственно $|\boldsymbol{a}|$, $|\boldsymbol{b}|$, $|\boldsymbol{c}|$, $|\boldsymbol{x}|$. Нулевой вектор обычно обозначают $\mathbf{0}$.

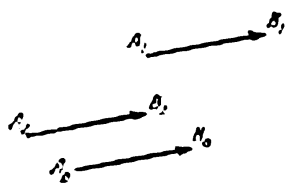


Рис. 7.18

Каждый ненулевой свободный вектор $\boldsymbol{a}$ имеет бесконечно много изображений в виде направленных отрезков, отложенных от разных точек плоскости (пространства) $A_1, A_2, \ldots$ $\ldots, A_n, \ldots$ так, что длины отрезков A_1B_1, $A_2B_2, \ldots, A_nB_n, \ldots$ равны длине вектора $\boldsymbol{a}$, а лучи A_1B_1, $A_2B_2, \ldots, A_nB_n, \ldots$ сонаправлены, и их направление совпадает с направлением вектора $\boldsymbol{a}$ (рис. 7.18). Нулевой свободный вектор $\mathbf{0}$ также имеет бесконечно много изображений в виде точек плоскости.

Свободный вектор обычно называют просто *вектором*. Множество равных между собой векторов, принадлежащих одной прямой, называется *скользящим* вектором. Наряду со свободными и скользящими векторами рассматриваются *связанные* векторы, которые характеризуются модулем, направлением и положением начальной точки, называемой *точкой приложения*. Два связанных вектора считаются равными, если они имеют не только равные модули и одинаковые направления, но и общую начальную точку, т. е. равные связанные векторы — это совпадающие векторы.

Вектор на плоскости может быть введен с использованием понятия изометрии или, точнее говоря, частного случая изометрии — понятия параллельного переноса.

Изометрией называется отображение плоскости на себя, при котором сохраняются расстояния между точками.

Параллельным переносом называется отображение плоскости на себя, при котором все точки плоскости перемещаются в одном и том же направлении на одно и то же расстояние.

В этих терминах *вектором* (параллельным переносом), определяемым упорядоченной парой несовпадающих точек $(A; B)$, называется преобразование плоскости, при котором каждая точка M плоскости отображается на такую точку M_1 этой плоскости, что луч MM_1 сонаправлен с лучом AB и расстояние $|MM_1|$ равно расстоянию $|AB|$.

Любая пара совпадающих точек задает тождественное преобразование (т. е. преобразование, которое любую точку плоскости отображает на себя). Тождественное преобразование называют *нулевым вектором* и обозначают символом $\mathbf{0}$.

Сложение векторов. Пусть $\boldsymbol{a}$ и $\boldsymbol{b}$ — два ненулевых вектора. Отложим вектор $\boldsymbol{a}$ от точки O и обозначим его конец буквой A (рис. 7.19): $\overrightarrow{OA}=\boldsymbol{a}$. Отложим от точки A вектор $\boldsymbol{b}$ и обозначим его конец буквой B: $\overrightarrow{AB}=\boldsymbol{b}$. Вектор с началом в точке O и концом в точке B ($\overrightarrow{OB}=\boldsymbol{c}$) называют *суммой* векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ и пишут

$$\boldsymbol{c}=\boldsymbol{a}+\boldsymbol{b}.$$

О векторе $\boldsymbol{c}$ говорят, что он получен в результате сложения векторов $\boldsymbol{a}$ и $\boldsymbol{b}$.

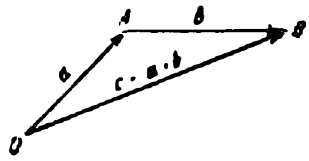


Рис. 7.19

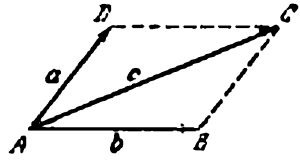


Рис. 7.20

Свойства операции сложения векторов:

1) свойство нулевого вектора:

$$\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{a};$$

2) ассоциативность:

$$(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c});$$

3) коммутативность:

$$\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}.$$

Сумму двух неколлинеарных векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ можно построить по правилу, называемому *правилом параллелограмма*. Отложим от точки A векторы $\overrightarrow{AB}=\boldsymbol{b}$ и $\overrightarrow{AD}=\boldsymbol{a}$ (рис. 7.20). Через конец вектора $\boldsymbol{a}$ проведем прямую, параллельную вектору $\boldsymbol{b}$, а через конец вектора $\boldsymbol{b}$ — прямую, параллельную вектору $\boldsymbol{a}$. Пусть C — точка пересечения построенных прямых. Вектор $\overrightarrow{AC}=\boldsymbol{c}$ — сумма векторов $\boldsymbol{a}$ и $\boldsymbol{b}$:

$$\boldsymbol{c}=\boldsymbol{a}+\boldsymbol{b}.$$

Вектором, *противоположным* вектору $\boldsymbol{a}$, называют вектор, обозначаемый $-\boldsymbol{a}$, такой, что сумма векторов $\boldsymbol{a}$ и $-\boldsymbol{a}$ равна нулевому вектору:

$$\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}.$$

Вектор, противоположный вектору $\overrightarrow{AB}$, обозначается также $\overrightarrow{BA}$:

$$\overrightarrow{AB}+\overrightarrow{BA}=\overrightarrow{AA}=\boldsymbol{0}.$$

Ненулевые противоположные векторы имеют равные длины ($|\boldsymbol{a}|=|-\boldsymbol{a}|$) и противоположные направления.

Разность двух векторов есть сумма вектора $\boldsymbol{a}$ и вектора, противоположного вектору $\boldsymbol{b}$ (рис. 7.21), т. е. вектор

$$\boldsymbol{a}+(-\boldsymbol{b}).$$

Разность векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ обозначают $\boldsymbol{a}-\boldsymbol{b}$.

Разность двух векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ может быть получена с помощью *правила треугольника*. Отложим от точки A вектор $\boldsymbol{a}$ (рис. 7.22): $\overrightarrow{AB}=\boldsymbol{a}$; от конца вектора $\overrightarrow{AB}$ отложим вектор $\overrightarrow{BC}=-\boldsymbol{b}$; вектор $\overrightarrow{AC}=\boldsymbol{c}$ — разность векторов $\boldsymbol{a}$ и $\boldsymbol{b}$:

$$\boldsymbol{c}=\boldsymbol{a}-\boldsymbol{b}.$$

Умножение вектора на число. *Произведением* ненулевого вектора $\boldsymbol{a}$ на число $\lambda\neq 0$ называют вектор, имеющий направление вектора $\boldsymbol{a}$, если λ положительно, и противоположное

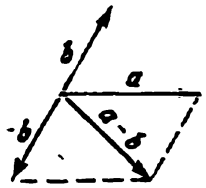


Рис. 7.21

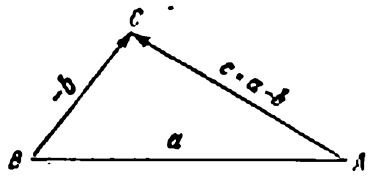


Рис. 7.22

направление, если λ отрицательно; длина этого вектора равна произведению длины вектора $\boldsymbol{a}$ на абсолютное значение (модуль) числа λ. Произведение вектора $\boldsymbol{a}$ на число λ обозначается через $\lambda\boldsymbol{a}$.

В случае, если $\boldsymbol{a}=0$ или $\lambda=0$, соответственно считают

$$\lambda\cdot 0=0 \text{ для любого } \lambda, \qquad 0\cdot\boldsymbol{a}=0 \text{ для любого } \boldsymbol{a}.$$

Свойства операции умножения вектора на число:

1) коммутативность:

$$\lambda\boldsymbol{a}=\boldsymbol{a}\lambda;$$

2) ассоциативность:

$$\lambda(\mu\boldsymbol{a})=(\lambda\mu)\,\boldsymbol{a};$$

3) дистрибутивность относительно сложения векторов:

$$\lambda(\boldsymbol{a}+\boldsymbol{b})=\lambda\boldsymbol{a}+\lambda\boldsymbol{b};$$

4) дистрибутивность относительно сложения чисел:

$$(\lambda+\mu)\,\boldsymbol{a}=\lambda\boldsymbol{a}+\mu\boldsymbol{a}.$$

Два ненулевых вектора называют *коллинеарными*, если их направления совпадают или противоположны.

Нулевой вектор по определению считается коллинеарным любому вектору.

Необходимым и достаточным условием коллинеарности ненулевых векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ является существование числа λ, удовлетворяющего равенству $\boldsymbol{b}=\lambda\boldsymbol{a}$.

Необходимым и достаточным условием коллинеарности векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ на плоскости, заданных своими координатами относительно прямоугольной декартовой системы координат Oxy:

$$\boldsymbol{a}=(x_1;\ y_1), \qquad \boldsymbol{b}=(x_2;\ y_2),$$

является равенство нулю определителя второго порядка:

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}=0.$$

Необходимым и достаточным условием коллинеарности двух векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ в пространстве, заданных своими координатами относительно прямоугольной декартовой системы координат $Oxyz$:

$$\boldsymbol{a}=(x_1;\ y_1;\ z_1), \qquad \boldsymbol{b}=(x_2;\ y_2;\ z_2),$$

является равенство нулю всех следующих определителей второго порядка:

$$\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}, \quad \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}, \quad \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}.$$

Три ненулевых вектора называются *компланарными*, если лучи, задающие их направления, принадлежат прямым, параллельным некоторой плоскости. Если среди трех векторов имеется хотя бы один нулевой, то такие векторы также считаются компланарными.

Необходимым и достаточным условием компланарности трех векторов $\boldsymbol{a}$, $\boldsymbol{b}$ и $\boldsymbol{c}$, заданных своими координатами в прямоугольной декартовой системе координат $Oxyz$:

$$\boldsymbol{a}=(x_1;\ y_1;\ z_1), \qquad \boldsymbol{b}=(x_2;\ y_2;\ z_2),$$
$$\boldsymbol{c}=(x_3;\ y_3;\ z_3),$$

является равенство нулю определителя третьего порядка:

$$\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}=0.$$

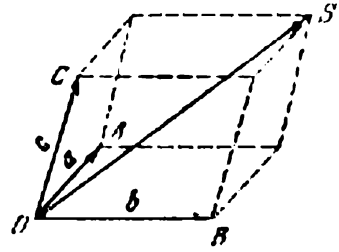


Рис. 7.23

Если векторы $\boldsymbol{a}$ и $\boldsymbol{b}$ не коллинеарны, то любой вектор $\boldsymbol{c}$, компланарный с векторами $\boldsymbol{a}$ и $\boldsymbol{b}$, можно представить единственным образом в виде

$$\boldsymbol{c}=x\boldsymbol{a}+y\boldsymbol{b}, \tag{1}$$

т. е. существует единственная упорядоченная пара чисел $(x;\ y)$, для которой выполняется равенство (1).

Представление вектора $\boldsymbol{c}$, компланарного с векторами $\boldsymbol{a}$ и $\boldsymbol{b}$, в виде суммы $\boldsymbol{c}=x\boldsymbol{a}+y\boldsymbol{b}$ называется *разложением* вектора $\boldsymbol{c}$ по векторам $\boldsymbol{a}$ и $\boldsymbol{b}$.

Правило параллелепипеда. Для нахождения вектора, являющегося суммой трех некомпланарных векторов $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, часто используют правило, называемое *правилом параллелепипеда*.

От произвольной точки пространства O откладывают векторы $\overrightarrow{OA}=\boldsymbol{a}$, $\overrightarrow{OB}=\boldsymbol{b}$ и $\overrightarrow{OC}=\boldsymbol{c}$. Строят параллелепипед так, чтобы отрезки OA, OB, OC были его ребрами (рис. 7.23). Вектор $\overrightarrow{OS}$, где OS — диагональ параллелепипеда,— сумма трех данных векторов:

$$\overrightarrow{OS}=\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}.$$

Разложение вектора по трем некомпланарным векторам. Пусть $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ — три некомпланарных ненулевых вектора. Тогда любой пространственный вектор $\boldsymbol{d}$ единственным образом может быть представлен в виде суммы:

$$\boldsymbol{d}=x\boldsymbol{a}+y\boldsymbol{b}+z\boldsymbol{c}, \tag{2}$$

т. е. существует единственная упорядоченная тройка чисел $(x;\ y;\ z)$, для которой выполняется равенство (2).

3.2. Угол между векторами. Скалярное произведение векторов. *Углом* между двумя ненулевыми векторами называется угол между направлениями этих векторов. Угол между двумя векторами $\boldsymbol{a}$ и $\boldsymbol{b}$ часто обозначается $(\widehat{\boldsymbol{a}, \boldsymbol{b}})$.

Если угол между векторами $\boldsymbol{a}$ и $\boldsymbol{b}$ равен 90°, то эти векторы называют *перпендикулярными* и пишут $\boldsymbol{a} \perp \boldsymbol{b}$.

Скалярным произведением двух ненулевых векторов $\boldsymbol{a}$ и $\boldsymbol{b}$, обозначаемым $(\boldsymbol{a}, \boldsymbol{b})$ (или $\boldsymbol{a} \cdot \boldsymbol{b}$), называется число, равное произведению длин векторов на косинус угла между ними:

$$(\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \cos \varphi,$$

где $\varphi = (\widehat{\boldsymbol{a}, \boldsymbol{b}})$. Если хотя бы один из двух векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ нулевой, то скалярное произведение этих двух векторов считается равным нулю. О скалярном произведении $(\boldsymbol{a}, \boldsymbol{b})$ говорят, что оно получено в результате *скалярного умножения* вектора $\boldsymbol{a}$ на вектор $\boldsymbol{b}$.

Свойства скалярного умножения векторов:

1) коммутативность:

$$(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{a});$$

2) ассоциативность относительно умножения на число:

$$((k\boldsymbol{a}), \boldsymbol{b}) = k(\boldsymbol{a}, \boldsymbol{b}).$$

Операции сложения и скалярного умножения векторов связаны дистрибутивным законом умножения относительно сложения:

$$(\boldsymbol{a}, (\boldsymbol{b} + \boldsymbol{c})) = (\boldsymbol{a}, \boldsymbol{b}) + (\boldsymbol{a}, \boldsymbol{c}).$$

Необходимым и достаточным условием перпендикулярности двух ненулевых векторов является равенство нулю их скалярного произведения.

Скалярное произведение вектора на тот же вектор равно квадрату числового значения его длины. Произведение $\boldsymbol{a} \cdot \boldsymbol{a}$ называется *скалярным квадратом* вектора $\boldsymbol{a}$ и обозначается $(\boldsymbol{a})^2$.

3.3. Координаты вектора на плоскости. Пусть Oxy — прямоугольная декартова система координат с базисными векторами $(\boldsymbol{i}; \boldsymbol{j})$. Любой вектор $\boldsymbol{a}$, принадлежащий плоскости, может быть разложен по векторам базиса $\boldsymbol{i}, \boldsymbol{j}$, т. е. для любого вектора $\boldsymbol{a}$ существует, и притом только одна, упорядоченная пара чисел $(x_0; y_0)$ такая, что

$$\boldsymbol{a} = x_0\boldsymbol{i} + y_0\boldsymbol{j}.$$

Числа x_0, y_0 называют *координатами* вектора $\boldsymbol{a}$ в базисе $(\boldsymbol{i}; \boldsymbol{j})$ и пишут

$$\boldsymbol{a} = (x_0; y_0).$$

Если в прямоугольной системе координат точки A и B имеют координаты $(x_1; y_1)$ и $(x_2; y_2)$, то координатами вектора $\overrightarrow{AB}$ будет упорядоченная пара чисел $(x_2 - x_1; y_2 - y_1)$, т. е.

$$\overrightarrow{AB} = (x_2 - x_1; y_2 - y_1).$$

Правила действий над векторами, заданными координатами. Пусть в базисе $(\boldsymbol{i}; \boldsymbol{j})$ векторы $\boldsymbol{a}$ и $\boldsymbol{b}$ заданы своими координатами:

$$\boldsymbol{a} = (x_1; y_1), \qquad \boldsymbol{b} = (x_2; y_2).$$

1. Координаты суммы двух векторов равны сумме соответствующих координат слагаемых:

$$a + b = (x_1 + x_2;\ y_1 + y_2).$$

2. Координаты разности двух векторов равны разности соответствующих координат этих векторов:

$$a - b = (x_1 - x_2;\ y_1 - y_2).$$

3. Координаты произведения вектора на число равны произведению соответствующих координат данного вектора на это число:

$$pa = (px_1;\ py_1).$$

4. Скалярное произведение двух векторов равно сумме произведений соответствующих координат этих векторов:

$$(a,\ b) = x_1x_2 + y_1y_2.$$

Модуль вектора a, заданного своими координатами $(x_1;\ y_1)$, вычисляется по формуле

$$|a| = \sqrt{x_1^2 + y_1^2}.$$

Угол φ между двумя ненулевыми векторами a и b, заданными своими координатами $(x_1;\ y_1)$ и $(x_2;\ y_2)$ соответственно, вычисляется из равенства

$$\cos\varphi = \cos(\widehat{a,\ b}) = \frac{x_1x_2 + y_1y_2}{\sqrt{x_1^2 + y_1^2}\sqrt{x_2^2 + y_2^2}}.$$

Косинусы углов между вектором $a = (x_1;\ x_2)$ и векторами базиса i, j называются *направляющими косинусами* вектора a и вычисляются по формулам

$$\cos\alpha = \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \quad \cos\beta = \frac{x_2}{\sqrt{x_1^2 + x_2^2}}.$$

Направляющие косинусы любого вектора a связаны равенством

$$\cos^2\alpha + \cos^2\beta = 1.$$

3.4. Координаты вектора в пространстве. Пусть $Oxyz$ — прямоугольная декартова система координат с тройкой базисных векторов $(i;\ j;\ k)$. Любой пространственный вектор a может быть разложен по векторам базиса i, j, k, т. е. для любого вектора a существует, и притом только одна, упорядоченная тройка чисел $(x_0;\ y_0;\ z_0)$ такая, что

$$a = x_0i + y_0j + z_0k.$$

Числа x_0, y_0, z_0 называют *координатами* вектора a в базисе $(i;\ j;\ k)$ и пишут

$$a = (x_0;\ y_0;\ z_0).$$

Если в прямоугольной системе координат точки A и B имеют координаты $(x_1;\ y_1;\ z_1)$ и $(x_2;\ y_2;\ z_2)$, то координатами вектора $\overrightarrow{AB}$ будет упорядоченная тройка чисел $(x_2 - x_1;\ y_2 - y_1;\ z_2 - z_1)$, т. е.

$$\overrightarrow{AB} = (x_2 - x_1;\ y_2 - y_1;\ z_2 - z_1).$$

Правила действий над векторами, заданными своими координатами. Пусть в базисе $(i;\ j;\ k)$ векторы a и

b заданы своими координатами:

$$a=(x_1;\ y_1;\ z_1),\quad b=(x_2;\ y_2;\ z_2).$$

1. Координаты суммы двух векторов равны сумме соответствующих координат слагаемых:

$$a+b=(x_1+x_2;\ y_1+y_2;\ z_1+z_2).$$

2. Координаты разности двух векторов равны разности соответствующих координат этих векторов:

$$a-b=(x_1-x_2;\ y_1-y_2;\ z_1-z_2).$$

3. Координаты произведения вектора на число равны произведению соответствующих координат данного вектора на это число:

$$pa=(px_1;\ py_1;\ pz_1).$$

4. Скалярное произведение двух векторов равно сумме произведений соответствующих координат этих векторов:

$$(a,\ b)=x_1x_2+y_1y_2+z_1z_2.$$

Модуль вектора a, заданного своими координатами $(x_1;\ y_1;\ z_1)$ относительно некоторого прямоугольного базиса $(i;\ j;\ k)$, вычисляется по формуле

$$|a|=\sqrt{x_1^2+y_1^2+z_1^2}.$$

Угол φ между двумя ненулевыми пространственными векторами a и b, задаваемыми своими координатами $(x_1;\ y_1;\ z_1)$ и $(x_2;\ y_2;\ z_2)$, вычисляется из равенства

$$\cos\varphi=\cos\widehat{(a,\ b)}=\frac{x_1x_2+y_1y_2+z_1z_2}{\sqrt{x_1^2+y_1^2+z_1^2}\sqrt{x_2^2+y_2^2+z_2^2}}.$$

Косинусы углов между вектором $a=(x_1;\ y_1;\ z_1)$ и векторами базиса i, j, k называются *направляющими косинусами* вектора a и вычисляются по формулам

$$\cos\alpha=\frac{x_1}{\sqrt{x_1^2+y_1^2+z_1^2}},\quad \cos\beta=\frac{y_1}{\sqrt{x_1^2+y_1^2+z_1^2}},$$

$$\cos\gamma=\frac{z_1}{\sqrt{x_1^2+y_1^2+z_1^2}}.$$

Направляющие косинусы любого ненулевого вектора a связаны равенством

$$\cos^2\alpha+\cos^2\beta+\cos^2\gamma=1.$$

3.5. **Векторное произведение.** *Векторным произведением* двух векторов a и b, обозначаемым $[a,\ b]$ (или $a\times b$), называется вектор, определяемый следующими тремя условиями:

1) модуль вектора $[a,\ b]$ равен $|a|\cdot|b|\sin\varphi$, где φ — угол между векторами a и b;

2) вектор $[a,\ b]$ перпендикулярен как вектору a, так и вектору b;

3) упорядоченная тройка векторов $(a;\ b;\ [a,\ b])$, отложенных от одной точки, образует правый (в общем случае косоугольный) базис (рис. 7.24).

Если хотя бы один из векторов a или b нулевой, то векторное произведение двух векторов считается равным нулевому вектору:

$$[a, b] = 0.$$

Векторное произведение двух ненулевых векторов a и b, заданных своими координатами: $a = (x_1; y_1; z_1)$ и $b = (x_2; y_2; z_2)$, есть вектор

$$[a, b] = \left(\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}; \quad \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}; \quad \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \right),$$

где $\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}$, $\begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}$, $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$ — определители второго порядка.

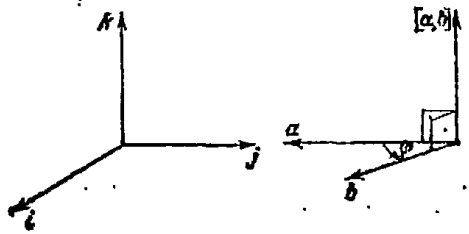


Рис. 7.24

Свойства векторного произведения:

$$[a, b] = -[b, a];$$
$$[(pa), b] = p[a, b] \quad (p \text{ — число});$$
$$[(a + b), c] = [a, c] + [b, c];$$
$$[a, [b, c]] = b(a, c) - c(a, b);$$
$$[[a, b], c] = b(a, c) - a(b, c).$$

3.6. Смешанное (скалярно-векторное) произведение векторов. Пусть $Oxyz$ — прямоугольная декартова система координат с правым базисом

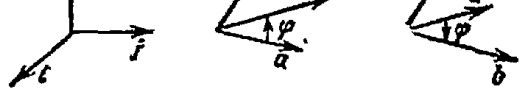


Рис. 7.25

$(i; j; k)$. *Смешанным произведением* трех ненулевых некомпланарных векторов, заданных своими координатами относительно базиса $(i; j; k)$

$$a = (x_1; y_1; z_1), \quad b = (x_2; y_2; z_2), \quad c = (x_3; y_3; z_3),$$

называется число, абсолютная величина которого равна объему параллелепипеда с ребрами a, b, c и которое положительно, если упорядоченная тройка векторов $(a; b; c)$ образует правый базис, и отрицательно, если левый *) (рис. 7.25).

*) В случае, если базис (i, j, k) левый, смешанное произведение векторов a, b, c положительно, если тройка векторов $(a; b; c)$ образует левый базис, и отрицательно в противном случае.

Смешанное произведение векторов $(a; b; c)$ часто обозначается $\langle a, b, c\rangle$.

Если векторы a, b, c компланарны, то смешанное произведение считается равным нулю:

$$\langle a, b, c\rangle = 0.$$

Смешанное произведение векторов a, b, c, заданных своими координатами, вычисляется как определитель третьего порядка:

$$\langle a, b, c\rangle = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

Свойства смешанного произведения:

$$\langle a, b, c\rangle = ([a, b], c);$$

$$\langle a, b\ c\rangle = ([b, c], a) = (c, [a, b]) = -(b, [a, c]).$$

Глава 8

ТЕОРИЯ ПРЕДЕЛОВ

Интуитивное понятие о предельном переходе использовалось еще учеными Древней Греции при вычислении площадей и объемов различных геометрических фигур. Методы решения таких задач в основном были развиты Архимедом. При создании дифференциального и интегрального исчислений математики XVII века (и прежде всего И. Ньютон) также явно или неявно использовали понятие предельного перехода. Впервые определение понятия предела было введено в работе Дж. Валлиса «Арифметика бесконечных величин» (XVII век). В XIX веке в работах О. Коши теория пределов была использована для строгого обоснования математического анализа. Дальнейшей разработкой теории пределов занимались К. Вейерштрасс и Б. Больцано. С помощью теории пределов во второй половине XIX века было, в частности, обосновано использование в анализе бесконечных рядов, которые явились удобным аппаратом для построения новых функций.

§ 1. Числовые последовательности

1.1. Понятие числовой последовательности. Рассмотрим множество чисел

$$1,\ \frac{1}{2},\ \frac{1}{3},\ \frac{1}{4},\ \ldots,\ \frac{1}{n},\ \ldots \qquad (1)$$

Будем считать, что каждое число этого множества снабжено номером, соответствующим тому месту, которое оно занимает в записи (1). Так, будем считать, что число 1, стоящее на первом месте, имеет номер 1; число 1/2, стоящее на втором месте,— номер 2; число 1/3 — номер 3 и т. д. И наоборот, в соответствии с этим соглашением о нумерации множества чисел (1), какой бы номер мы ни указали, в множестве чисел (1) найдется единственное число, имеющее указанный номер. Например, номер 4 имеет число 1/4, номер 10 — число 1/10. Таким образом, в приведенном множестве чисел (1) каждое число имеет вполне определенный номер и полностью определяется им.

Множество чисел, занумерованных либо конечным отрезком натурального ряда, либо всеми натуральными числами, называют *числовой последовательностью.*

В первом случае последовательность называется *конечной*, во втором — *бесконечной.*

Элементы этого числового множества называются *членами* последовательности и обычно обозначаются так: первый член a_1, второй a_2, ..., n-й — a_n и т. д. Вся числовая последовательность обозначается

$$a_1,\ a_2,\ a_3,\ \ldots,\ a_n,\ \ldots, \quad \text{или} \quad (a_n).$$

Числа, образующие числовую последовательность, не обязательно должны быть все различны. Бесконечная числовая последовательность

может состоять и из конечного числа различных чисел (и даже из одного числа). Например, из двух чисел — нуля и единицы — можно составить бесконечную числовую последовательность

$$0,\ 1,\ 0,\ 1,\ \ldots,$$

располагая нули на нечетных местах, а единицы — на четных.

Понятие числовой последовательности может быть введено также с помощью понятия отображения одного числового множества на другое.

Пусть $\mathbf{N}$ — множество всех натуральных чисел, A — некоторое конечное или счетное числовое множество. Зададим отображение f множества всех натуральных чисел $\mathbf{N}$ на множество A, т. е. укажем правило, по которому для каждого натурального числа $n \in \mathbf{N}$ выбирается некоторый вполне определенный элемент числового множества A, который обозначим a_n (индекс n указывает на то, что число a_n соответствует n), и, наоборот, каждому элементу a_n будет соответствовать хотя бы одно натуральное число n.

Отображение множества $\mathbf{N}$ на некоторое конечное или счетное числовое множество A называется *бесконечной числовой последовательностью*; число a_n называется *членом* последовательности, а индекс n — его *номером*.

Задать бесконечную числовую последовательность — это значит задать отображение, при котором каждому натуральному числу n соответствует один и только один элемент числового множества A.

Иногда бесконечную числовую последовательность вводят, используя понятие функции: *бесконечной числовой последовательностью* называется числовая функция, определенная на множестве всех натуральных чисел.

Суммой двух последовательностей (x_n) и (y_n) называют последовательность (s_n), членами которой являются суммы членов последовательностей (x_n) и (y_n) с одинаковыми номерами:

$$s_n = x_n + y_n, \quad n \in \mathbf{N}.$$

Произведением, разностью и *частным* двух последовательностей (x_n) и (y_n) называют соответственно последовательности (m_n), (q_n) и (r_n), члены которых вычисляются по правилам:

$$m_n = x_n \cdot y_n; \quad q_n = x_n - y_n, \quad r_n = x_n / y_n, \quad n \in \mathbf{N},$$

причем частное двух последовательностей (т. е. последовательность (r_n)) определено лишь для таких последовательностей (y_n), у которых ни один член не равен нулю.

1.2. Некоторые способы задания последовательностей.

1. Аналитический способ задания последовательности. Задать последовательность аналитически — это значит указать формулу, позволяющую по номеру члена последовательности однозначно определить этот член. Формула, позволяющая вычислить любой член последовательности по его номеру, называется *формулой общего члена* числовой последовательности.

Например, формулы общего члена

$$a_n = \frac{(-1)^{n-1}}{n}, \quad b_n = \sin n, \quad c_n = \frac{1 + (-1)^n}{n}$$

задают соответственно следующие числовые последовательности:

$$1,\ -\frac{1}{2},\ \frac{1}{3},\ -\frac{1}{4},\ \frac{1}{5},\ \ldots,$$

$$\sin 1,\ \sin 2,\ \sin 3,\ \sin 4,\ \ldots,$$

$$0,\ 1,\ 0,\ \frac{1}{2},\ 0,\ \frac{1}{3},\ \ldots$$

2. Рекуррентный способ задания последовательности. Рекуррентный способ состоит в том, что задается первый член (или несколько первых членов) последовательности и указывается формула вычисления последующих членов последовательности по заданному первому члену (или нескольким первым членам).

Примеры рекуррентного задания последовательностей:

а) ряд натуральных чисел может быть задан следующим рекуррентным способом:

$$a_1 = 1, \quad a_{n+1} = a_n + 1, \quad n \in \mathbf{N};$$

б) последовательность чисел Фибоначчи (см. § 1 гл. 10) задается первыми двумя членами и формулой вычисления каждого следующего члена, начиная с третьего:

$$b_1 = 0, \quad b_2 = 1, \quad b_{n+2} = b_n + b_{n+1}, \quad n \in \mathbf{N}.$$

3. Последовательность может быть также задана описанием способа получения ее членов. Например, говорят, что последовательность

$$3{,}1; \quad 3{,}14; \quad 3{,}141; \quad 3{,}1415; \quad \ldots$$

состоит из приближенных значений числа π с недостатком с точностью до 0,1; 0,01; 0,001; 0,0001 и т. д. В подобных случаях, как правило, нельзя указать ни формулы общего члена последовательности, ни рекуррентного способа вычисления ее членов.

1.3. Геометрическое изображение членов последовательности. Под геометрическим изображением членов последовательности обычно пони-

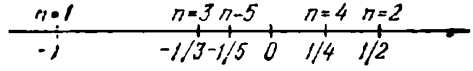


Рис. 8.1

мается изображение членов последовательности как точек числовой прямой (координатной оси). Например, первые пять членов последовательности, задаваемой формулой общего члена

$$a_n = \frac{(-1)^n}{n}, \quad n \in \mathbf{N}, \qquad (2)$$

изображены на рис. 8.1. Реже члены последовательности, рассматриваемой как функция натурального аргумента, изображаются точками в прямоугольной декартовой системе координат Oxy. На рис. 8.2 в прямоугольной декартовой системе координат изображены первые пять членов последовательности (2).

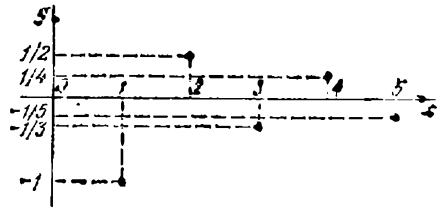


Рис. 8.2

1.4. Ограниченные последовательности. Последовательность (x_n) называется *ограниченной*, если существуют такие два числа m и M, что для всех $n \in \mathbf{N}$ выполняется неравенство

$$m \leqslant x_n \leqslant M. \qquad (3)$$

Последовательность (x_n) называется *ограниченной сверху*, если существует такое число M, что для всех $n \in \mathbf{N}$ выполняется неравенство $x_n \leqslant M$.

Последовательность (x_n) называется *ограниченной снизу*, если существует такое число m, что для всех $n \in \mathbf{N}$ выполняется неравенство

$$x_n \geqslant m.$$

Доказательство ограниченности последовательности сводится к нахождению таких чисел m и M, для которых неравенство (3) справедливо для всех $n \in \mathbf{N}$.

Примеры. 1. Последовательность (x_n), задаваемая формулой общего члена $x_n = n$, ограничена снизу (например, числом 0) и не ограничена сверху.

2. Последовательность (x_n), задаваемая формулой общего члена $x_n = -n^2/(n+1)$, ограничена сверху (например, числом $-1/2$) и не ограничена снизу.

3. Последовательности, задаваемые формулами общего члена $x_n = \sin n$, $y_n = (-1)^n/n$, ограничены числами -1; 1 и -1; 1/2 соответственно.

1.5. Монотонные последовательности. Последовательность (x_n) называется *возрастающей*, если каждый ее член, начиная со второго, больше предыдущего, т. е. если для любого натурального n выполняется неравенство $x_{n+1} > x_n$.

Последовательность (x_n) называется *убывающей*, если каждый член, начиная со второго, меньше предыдущего, т. е. если для любого натурального n выполняется неравенство $x_{n+1} < x_n$.

Последовательность (x_n) называется *неубывающей*, если каждый ее член, начиная со второго, не меньше предыдущего, т. е. если для любого натурального n выполняется неравенство $x_{n+1} \geqslant x_n$.

Последовательность (x_n) называется *невозрастающей*, если каждый ее член, начиная со второго, не больше предыдущего, т. е. если для любого натурального n выполняется неравенство $x_{n+1} \leqslant x_n$.

Возрастающие, убывающие, невозрастающие и неубывающие последовательности образуют класс *монотонных последовательностей* *).

Возрастающие и убывающие последовательности называют *строго монотонными*.

Доказательство монотонности последовательности состоит в доказательстве какого-либо из перечисленных выше неравенств для всех $n \in \mathbf{N}$.

Примеры 1. Доказать монотонность последовательности, задаваемой формулой общего члена

$$x_n = \frac{3n+1}{2n-1}, \quad n \in \mathbf{N}.$$

Рассмотрим разность между $n+1$-м и n-м членами:

$$x_{n+1} - x_n = \frac{3n+4}{2n+1} - \frac{3n+1}{2n-1} = \frac{-5}{(2n+1)(2n-1)}.$$

Числитель дроби, стоящей справа, отрицателен, знаменатель положителен при любом натуральном n; следовательно, вся дробь отрицательна при любом $n \in \mathbf{N}$, поэтому

$$x_{n+1} - x_n < 0 \Leftrightarrow x_{n+1} < x_n,$$

т. е. данная последовательность убывающая.

*) Во многих случаях (например, при исследовании сходимости последовательности) важна монотонность последовательности, начиная с некоторого (не обязательно первого) члена. Поэтому удобным (и более общим) оказывается такое определение *монотонной последовательности*: последовательность (x_n) называется *монотонно возрастающей* (*убывающей*), если найдется такой номер N_0, что для всех членов последовательности с номерами, бо́льшими N_0, выполняется неравенство $x_{n+1} > x_n$ ($x_{n+1} < x_n$). Такие последовательности — монотонные, начиная с некоторого номера.

2. Исследовать на монотонность последовательность, задаваемую формулой общего члена

$$x_n = n^2 - 9n - 100.$$

Вычислим разность между $n+1$-м и n-м членами:

$$x_{n+1} - x_n = [(n+1)^2 - 9(n+1) - 100] - [n^2 - 9n - 100] = 2n - 8.$$

Выражение, стоящее справа, отрицательно при $n = 1, 2, 3$, равно нулю при $n = 4$ и положительно при остальных значениях $n \in \mathbb{N}$.

По определению монотонной последовательности требуется, чтобы сразу для всех n выполнялось одно из четырех неравенств: $x_{n+1} > x_n$, $x_{n+1} \geqslant x_n$, $x_{n+1} < x_n$, $x_{n+1} \leqslant x_n$, но для данной последовательности ни одно из этих неравенств сразу для всех значений $n \in \mathbb{N}$ не выполняется, и поэтому рассматриваемая последовательность будет немонотонной. (Эта последовательность — монотонно возрастающая, начиная с пятого члена.)

§ 2. Предел последовательности

2.1. Понятие предела последовательности. Рассмотрим бесконечную числовую последовательность

$$1 - \frac{1}{1}, \quad 1 - \frac{1}{2}, \quad 1 - \frac{1}{3}, \quad 1 - \frac{1}{4}, \ldots,$$

задаваемую формулой общего члена

$$x_n = 1 - \frac{1}{n}, \quad n \in \mathbb{N}. \tag{1}$$

Нетрудно проверить, что эта последовательность ограничена (снизу числом 0, сверху числом 1) и монотонно возрастает. Если изображать члены последовательности точками числовой прямой, то все эти точки будут расположены между точками 0 и 1 (рис. 8.3). С ростом

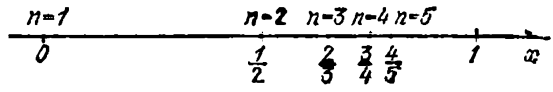


Рис. 8.3

номеров точки, изображающие члены последовательности, будут приближаться к точке числовой оси, имеющей абсциссу 1, и абсолютная величина разности

$$|x_n - 1| = \left|1 - \frac{1}{n} - 1\right| = \frac{1}{n}$$

неограниченно уменьшается с ростом n. Например, для всех членов последовательности, начиная с 11-го, абсолютная величина разности $x_n - 1$ меньше 0,1; для всех членов последовательности, начиная со 101-го, она меньше 0,01 и т. д. Вообще, какое бы малое положительное число ε мы ни выбрали, в данной последовательности найдется такой член (номер которого мы будем обозначать N_0), что для всех членов последовательности с номерами, большими N_0, будет выполняться неравенство

$$|x_n - 1| < \varepsilon,$$

т. е. для любого положительного числа ε существует номер N_0 (для каждого числа ε, вообще говоря, свой), начиная с которого все члены последовательности принадлежат промежутку $(1-\varepsilon;\ 1+\varepsilon)$.

Число 1 называют *пределом последовательности* (x_n), задаваемой формулой общего члена (1), и пишут *)

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)=1.$$

Пусть (x_n) — некоторая бесконечная числовая последовательность. Число a называют *пределом последовательности* (x_n) и пишут

$$\lim_{n\to\infty} x_n = a,$$

если для любого сколь угодно малого положительного числа ε существует такой номер N_0, что все члены последовательности (x_n) с номерами, большими N_0, удовлетворяют неравенству

$$|x_n - a| < \varepsilon. \tag{2}$$

Если число a является пределом последовательности (x_n), то о последовательности (x_n) говорят, что она *сходится* к числу a или имеет *пределом* число a. Последовательности, не имеющие конечного предела, иногда называют *расходящимися*.

Неравенство (2) может быть также записано в виде

$$a-\varepsilon < x_n < a+\varepsilon.$$

Промежуток $(a-\varepsilon;\ a+\varepsilon)$ называют *ε-окрестностью* точки a. Используя это понятие, можно дать и такое определение предела последовательности.

Число a называется *пределом* числовой последовательности (x_n), если для любого положительного числа ε существует такой номер N_0, что все члены последовательности, начиная с x_{N_0+1}, принадлежат ε-окрестности точки a. При этом необходимо иметь в виду, что номер N_0 зависит от выбора числа ε. Для того чтобы это подчеркнуть, часто вместо N_0 пишут $N_0(\varepsilon)$. При уменьшении ε соответствующий номер $N_0(\varepsilon)$, как правило, увеличивается. Исключение составляет случай, когда все члены последовательности (x_n) равны:

$$x_1 = x_2 = x_3 = \ldots = x_n = x_{n+1} = \ldots$$

Предел такой последовательности

$$a = \lim_{n\to\infty} x_n$$

будет равен членам этой последовательности. В этом случае неравенство (2) будет выполняться для любого $\varepsilon > 0$ сразу при всех значениях $n \in \mathbb{N}$.

Пример. Доказать, что предел числовой последовательности, задаваемой формулой общего члена

$$x_n = \frac{n}{n+1}, \quad n \in \mathbb{N},$$

равен 1.

Покажем, что для любого $\varepsilon > 0$ существует такой номер N_0, что $|x_n - 1| < \varepsilon$ для всех $n > N_0$. Так как

$$|x_n - 1| = \left|\frac{n}{n+1} - 1\right| = \left|-\frac{1}{n+1}\right| = \frac{1}{n+1}, \tag{3}$$

*) lim есть сокращение латинского слова «limes», означающего «предел».

то неравенство $|x_n - 1| < \varepsilon$ можно записать в виде

$$\frac{1}{n+1} < \varepsilon. \qquad (4)$$

Разрешая это неравенство относительно n, получаем

$$n > \frac{1-\varepsilon}{\varepsilon}. \qquad (5)$$

Из неравенства (5) следует, что если в качестве числа $N_0(\varepsilon)$ выбрать любое натуральное число *), большее чем $(1-\varepsilon)/\varepsilon$, то из неравенств (4), (5) и равенства (3) следует, что все члены последовательности с номерами $n > N_0(\varepsilon)$ будут удовлетворять неравенству

$$|x_n - 1| < \varepsilon.$$

По определению предела это и означает, что число 1 является пределом заданной последовательности.

Говорят, что последовательность (x_n) *стремится* к $+\infty$ $(-\infty)$, и пишут

$$x_n \longrightarrow +\infty \qquad (x_n \longrightarrow -\infty),$$

если для любого сколь угодно большого положительного числа E существует такой номер N_0 (зависящий от E), что все члены последовательности (x_n) с номерами, бо́льшими $N_0(E)$, удовлетворяют неравенству

$$x_n > E \qquad (x_n < -E).$$

Так, последовательности, задаваемые формулами общих членов

$$x_n = n^2 \quad \text{и} \quad x_n = -n,$$

стремятся соответственно к $+\infty$ и $-\infty$.

2.2. Необходимое условие сходимости последовательности. Пусть последовательность (x_n) сходится к числу a, т. е.

$$\lim_{n\to\infty} x_n = a.$$

По определению предела для любого $\varepsilon > 0$ все члены данной последовательности, кроме, быть может, конечного их числа, попадают в ε-окрестность точки a, т. е.

$$a - \varepsilon < x_n < a + \varepsilon.$$

Пусть число членов этой последовательности, которые не попали в окрестность точки a, равно N_0, и пусть какие-то члены лежат справа от $a+\varepsilon$, а какие-то — слева от $a-\varepsilon$. Правее $a+\varepsilon$ лежит конечное число точек, поэтому, перебрав их, мы можем выбрать максимальный член, который обозначим через M. Аналогично минимальный из членов последовательности, лежащих левее точки $a-\varepsilon$, обозначим через m. Может оказаться, что справа от $a+\varepsilon$ (или слева от $a-\varepsilon$) не окажется ни одного члена. В этом случае в качестве числа M можно взять число $a+\varepsilon$ и соответственно в качестве числа m — число $a-\varepsilon$.

Таким образом, если x_n — сходящаяся последовательность, то существуют такие два числа m и M, что для всех натуральных n

$$m \leqslant x_n \leqslant M;$$

следовательно, (x_n) — ограниченная последовательность.

*) При $(1-\varepsilon)/\varepsilon > 0$ можно, например, взять число $[(1-\varepsilon)/\varepsilon]+1$. При $(1-\varepsilon)/\varepsilon < 0$ можно взять любое натуральное число. (Символом $[(1-\varepsilon)/\varepsilon]$ обозначается целая часть числа $(1-\varepsilon)/\varepsilon$. *Целой частью* действительного числа a называется целое число p такое, что $p \leqslant a < p+1$.)

Рассмотрим последовательность, задаваемую формулой общего члена,

$$x_n = (-1)^n, \quad n \in \mathbf{N}.$$

Эта последовательность ограничена (снизу числом —1, сверху числом 1). Используя определение предела, докажем, что эта последовательность не имеет предела.

Каждый член последовательности, имеющий нечетный номер, — точка числовой оси с абсциссой —1, а члены с четными номерами — точка с абсциссой 1. Предположим, что предел данной последовательности существует и равен a. Если a — некоторое число, отличное от 1 и —1, то всегда можно найти ε-окрестность точки a, в которую не попадет ни один из членов данной последовательности. Следовательно, если данная последовательность имеет предел, то он должен быть равен либо —1, либо 1. Допустим, что предел равен —1. Тогда для любого $\varepsilon > 0$ все члены последовательности, начиная с некоторого, должны попадать в ε-окрестность точки —1. Возьмем $\varepsilon = 0{,}1$. В промежуток (—1,1; —0,9) попадают все члены последовательности, стоящие на нечетных местах, но не попадает ни один из членов, стоящих на четных местах, т. е. не существует номера $N_0(\varepsilon)$, начиная с которого все члены последовательности попадают в ε-окрестность точки —1, и, следовательно, число —1 не является пределом данной последовательности. Аналогично доказывается, что и число 1 не является пределом последовательности. Значит, последовательность

$$-1, \ 1, \ -1, \ 1, \ -1, \ 1, \ \ldots$$

не имеет предела.

Таким образом, из ограниченности последовательности, задаваемой формулой общего члена $x_n = (-1)^n$, не следует ее сходимость.

Разобранные примеры позволяют сформулировать необходимое условие сходимости бесконечной числовой последовательности.

Для того чтобы последовательность сходилась, необходимо, чтобы она была ограничена.

2.3. Теоремы о пределах последовательностей.

1. Последовательность не может сходиться к двум различным пределам (теорема о единственности предела).

2. Предел последовательности, все члены которой равны одной и той же величине, равен этой величине.

3. Если последовательность (x_n) сходится к числу a и $a > p$ $(a < q)$, то и все члены последовательности, начиная с некоторого, тоже будут больше p (меньше q).

4. Если последовательность (x_n) имеет предел, то она ограничена.

5. Если последовательности (x_n) и (y_n) имеют пределы (сходятся), то их сумма (разность) тоже имеет предел, причем предел суммы (разности) равен сумме (разности) пределов:

$$\lim_{n \to \infty} (x_n \pm y_n) = \lim_{n \to \infty} x_n \pm \lim_{n \to \infty} y_n.$$

6. Если последовательности (x_n) и (y_n) имеют пределы (сходятся), то их произведение также имеет предел, равный произведению пределов сомножителей:

$$\lim_{n \to \infty} (x_n \cdot y_n) = \lim_{n \to \infty} x_n \cdot \lim_{n \to \infty} y_n.$$

Из теорем 2 и 6 следует, что постоянный множитель можно выносить за знак предела:

$$\lim_{n\to\infty}(cx_n)=c\lim_{n\to\infty}x_n.$$

7. Если последовательности (x_n) и (y_n) имеют пределы, причем $y_n\neq 0$ для всех n, начиная с некоторого N_ε, то последовательность (x_n/y_n) также имеет предел, равный отношению пределов последовательностей (x_n) и (y_n):

$$\lim_{n\to\infty}\frac{x_n}{y_n}=\frac{\lim\limits_{n\to\infty}x_n}{\lim\limits_{n\to\infty}y_n}.$$

Покажем теперь, как при помощи сформулированных теорем нахождение предела последовательности сводится к нахождению пределов простейших последовательностей.

Пример. Вычислить предел последовательности (x_n), задаваемой формулой общего члена

$$x_n=\frac{2n^2}{n^2+n+1}.$$

Воспользуемся теоремами о пределах последовательностей. При этом сразу применить теорему о пределе частного мы не можем, так как последовательности в числителе и знаменателе не имеют конечных пределов. Преобразуем дробь: разделим числитель и знаменатель на n^2:

$$\frac{2n^2}{n^2+n+1}=\frac{2}{1+\frac{1}{n}+\frac{1}{n^2}}.$$

Далее последовательно применим теоремы 7 и 5:

$$\lim_{n\to\infty}\frac{2n^2}{n^2+n+1}=\lim_{n\to\infty}\frac{2}{1+\frac{1}{n}+\frac{1}{n^2}}=\frac{\lim\limits_{n\to\infty}2}{\lim\limits_{n\to\infty}1+\lim\limits_{n\to\infty}\frac{1}{n}+\lim\limits_{n\to\infty}\frac{1}{n^2}}.$$

Используя определение предела, нетрудно доказать, что

$$\lim_{n\to\infty}\frac{1}{n}=0 \quad\text{и}\quad \lim_{n\to\infty}\frac{1}{n^2}=0.$$

В результате для заданной последовательности получаем

$$\lim_{n\to\infty}\frac{2n^2}{n^2+n+1}=\frac{2}{1+0+0}=2.$$

Следует отметить, что при нахождении предела заданной последовательности мы не совсем «законно» пользовались теоремами 7, 5. Например, используя теорему о пределе частного, мы поставили знак равенства между выражениями

$$\lim_{n\to\infty}\frac{2}{1+\frac{1}{n}+\frac{1}{n^2}}, \qquad \frac{\lim\limits_{n\to\infty}2}{\lim\limits_{n\to\infty}\left(1+\frac{1}{n}+\frac{1}{n^2}\right)},$$

хотя к этому моменту еще не знали, что пределы числителя и знаменателя второй дроби существуют. Сходимость последовательности, стоящей в знаменателе второй дроби, доказывается позже, после такого же «незаконного» применения теоремы 5. Соединение знаком равенства этих выражений, в сущности, означает, что сначала мы предполагаем, что условия применяемых теорем выпол-

жена, а затем, возвращаясь обратно по цепочке равенств, обосновываем законность поставленного ранее знака равенства.

2.4. Достаточное условие сходимости последовательности (теорема Вейерштрасса). Ограниченность последовательности является необходимым условием ее сходимости — из сходимости последовательности следует ее ограниченность, однако ограниченность последовательности не является достаточным условием ее сходимости. Монотонность последовательности кажется вовсе не связанной с ее сходимостью — монотонная последовательность может быть как сходящейся, так и расходящейся. Например, последовательность, задаваемая формулой общего члена $x_n = n$, монотонная и расходящаяся, а последовательность $x_n = 1/n$ монотонная и сходящаяся.

С другой стороны, немонотонные последовательности могут быть как сходящимися, так и расходящимися. Так, последовательность $x_n = (-1)^n/n$ сходящаяся, а последовательность $x_n = (-1)^n n$ расходящаяся.

Однако последовательности, которые одновременно являются монотонными и ограниченными, обязательно сходятся. Это свойство последовательностей, иногда называемое достаточным условием сходимости, формулируется в виде следующей теоремы.

Если последовательность монотонна и ограничена, то она сходится.

Эта теорема — о существовании предела последовательности; она не дает метода нахождения этого предела. Однако в ряде случаев, зная, что предел последовательности существует, можно указать способы его вычисления. При этом часто пользуются тем, что для любой сходящейся (не обязательно монотонной) последовательности (x_n) справедливо равенство

$$\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} x_n. \tag{6}$$

Прежде чем перейти к примерам, заметим, что при вычислении пределов с помощью равенства (6) удобно иметь рекуррентную, а не аналитическую форму задания последовательности, а потому во всех рассматриваемых ниже примерах аналитически задаваемые последовательности будут, по мере необходимости, записываться в рекуррентной форме *).

Пример 1. Вычислить предел последовательности

$$x_n = q^n, \quad 0 < q < 1.$$

Последовательность $x_n = q^n$ монотонно убывает (для любого $n \in \mathbb{N}$ при $0 < q < 1$ выполняется неравенство $x_n > x_{n+1}$) и ограничена ($0 < q^n < 1$), поэтому по теореме Вейерштрасса имеет предел:

$$\lim_{n\to\infty} x_n = a. \tag{7}$$

Запишем исходную последовательность в рекуррентной форме:

$$x_{n+1} = qx_n, \tag{8}$$

на основании равенства (6) и формулы (8) имеем

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} \Rightarrow \lim_{n\to\infty} x_n = \lim_{n\to\infty} (qx_n).$$

Так как постоянный множитель можно вынести из-под знака предела, то используя обозначение (7), последнее равенство можно записать в виде

$$\lim_{n\to\infty} x_n = q \lim_{n\to\infty} x_n \Rightarrow a = qa \Rightarrow a(1-q) = 0.$$

*) Следует иметь в виду, что не любая последовательность, заданная аналитическим способом, может быть записана в рекуррентной форме.

Но $q \neq 1$, поэтому $a=0$. Таким образом, предел данной последовательности равен нулю.

Пример 2. Вычислить предел последовательности

$$x_n=\frac{3n+1}{2n-1}.$$

Эта последовательность монотонно убывает, так как

$$x_{n+1}-x_n=-\frac{5}{(2n+1)(2n-1)}<0$$

при всех $n\in\mathbb{N}$; следовательно, наибольшим будет первый член, равный 4. С другой стороны, нетрудно проверить, что $x_n>0$ при всех $n\in\mathbb{N}$. Значит, последовательность ограничена: $0<x_n<4$ при всех $n\in\mathbb{N}$.

Таким образом, данная последовательность удовлетворяет условиям теоремы Вейерштрасса и имеет предел:

$$\lim_{n\to\infty} x_n=a.$$

Будем вычислять его с помощью формулы (6), переходя от аналитического способа задания последовательности к рекуррентному. Для этого разрешим равенство

$$x_n=\frac{3n+1}{2n-1}$$

относительно n:

$$n=\frac{x_n+1}{2x_n-3}.$$

Подставив полученное выражение для n в выражение для $n+1$-го члена

$$x_{n+1}=\frac{3n+1}{2n+1},$$

получим рекуррентную форму задания исходной последовательности:

$$x_{n+1}=\frac{11x_n-9}{4x_n-1}.$$

Подставим в левую часть равенства (6) вместо x_{n+1} его выражение через x_n:

$$\lim_{n\to\infty}\frac{11x_n-9}{4x_n-1}=\lim_{n\to\infty}x_n. \tag{9}$$

Используя теорему о пределе частного и учитывая, что предел последовательности (x_n) был обозначен через a, из равенства (9) получаем уравнение для нахождения числа a:

$$\frac{11a-9}{4a-1}=a.$$

Единственное решение данного уравнения: $a=3/2$. Таким образом, предел данной последовательности равен 3/2.

Очевидно, что предел последовательности $x_n=\frac{3n+1}{2n-1}$ можно вычислить и без применения теоремы Вейерштрасса, используя сформулированные в п. 2.3 теоремы о пределах последовательностей.

Пример 3. Вычислить предел последовательности, задаваемой формулой общего члена

$$x_n=\underbrace{\sqrt{c+\sqrt{c+\ldots+\sqrt{c}}}}_{n\text{ корней}},\quad c>0. \tag{10}$$

Докажем, что последовательность имеет предел, и вычислим его с помощью равенства (6). Эта последовательность может быть задана рекуррентной

формулой:

$$x_{n+1}=\sqrt{c+x_n}. \tag{11}$$

Докажем, что данная последовательность монотонно возрастает, т. е. что $x_{n+1}>x_n$ для всех $n\in\mathbb{N}$. Используя рекуррентную формулу (11), это неравенство можно записать в виде $\sqrt{c+x_n}>x_n$, что эквивалентно (с учетом условия $x_n>0$) неравенству $x_n^2-x_n-c<0$. С помощью формулы общего члена (10) последнее неравенство записывается в виде

$$\underbrace{\sqrt{c+\sqrt{c+\ldots+\sqrt{c}}}}_{n-1\text{ корней}}-\underbrace{\sqrt{c+\sqrt{c+\ldots+\sqrt{c}}}}_{n\text{ корней}}-c<0\Leftrightarrow$$

$$\Leftrightarrow\underbrace{\sqrt{c+\sqrt{c+\ldots+\sqrt{c}}}}_{n-1\text{ корней}}<\underbrace{\sqrt{c+\sqrt{c+\ldots+\sqrt{b}}}}_{n\text{ корней}}.$$

Возведя обе части последнего неравенства $n-1$ раз в квадрат, получим неравенство $0<\sqrt{c}$, справедливость которого очевидна. Значит, исходное неравенство верно при всех значениях $c>0$ и данная последовательность возрастающая.

Последовательность (10) ограничена сверху, например числом $\sqrt{c}+1$. Действительно, $x_1=\sqrt{c}$ меньше этого числа. Если теперь допустить, что $x_n<\sqrt{c}+1$, то и для следующего члена последовательности x_{n+1} получаем

$$x_{n+1}<\sqrt{c+\sqrt{c}+1}<\sqrt{c+2\sqrt{c}+1}=\sqrt{c}+1;$$

и ограниченность последовательности сверху может быть доказана методом математической индукции. Ограниченность последовательности снизу следует из того, что $x_n>0$ при всех $n\in\mathbb{N}$.

Итак, данная последовательность возрастающая и ограниченная; следовательно, по теореме Вейерштрасса она имеет предел, который обозначим через a ($\lim\limits_{n\to\infty}x_n=a$). Для вычисления числа a используем рекуррентную формулу (11), записывая ее в виде

$$x_{n+1}^2=c+x_n.$$

В силу сходимости последовательности имеет место равенство

$$\lim_{n\to\infty}x_{n+1}^2=\lim_{n\to\infty}(c+x_n),$$

из которого следует, что величина a удовлетворяет квадратному уравнению

$$a^2=c+a.$$

Множество решений данного квадратного уравнения:

$$a=\frac{1+\sqrt{4c+1}}{2},\qquad a=\frac{1-\sqrt{4c+1}}{2}.$$

Число $\frac{1-\sqrt{4c+1}}{2}$ отрицательно при любом $c>0$. Но так как отрицательное число не может быть пределом последовательности с положительными членами, то пределом рассматриваемой последовательности является число

$$\frac{1+\sqrt{4c+1}}{2}.$$

Пример 4. Вычислить предел последовательности, задаваемой формулой общего члена

$$x_n=\frac{2^{n-1}}{n!},\qquad n\in\mathbb{N}, \tag{12}$$

Данная последовательность убывающая, начиная со второго члена, что следует из неравенства

$$x_{n+1}-x_n=\frac{2^n}{(n+1)!}-\frac{2^{n-1}}{n!}=\frac{2^{n-1}}{n!}\left(\frac{2}{n+1}-1\right)=\frac{2^{n-1}}{n!}\cdot\frac{1-n}{n+1}=\frac{2^{n-1}(1-n)}{(n+1)!}<0.$$

[Следовательно, наибольшими членами последовательности (12) являются первый и второй члены, равные единице, и $x_n \leqslant 1$ для всех $n \in \mathbb{N}$. Так как $x_n > 0$ при всех $n \in \mathbb{N}$, то эта последовательность ограничена снизу числом нуль. Значит, последовательность (12) монотонна и ограничена: $0 < x_n \leqslant 1$. По теореме Вейерштрасса эта последовательность имеет некоторый предел a ($\lim\limits_{n\to\infty} x_n = a$).

Найдем значение a. Члены последовательности (12) связаны рекуррентной формулой

$$x_{n+1}=\frac{2^n}{(n+1)!}=\frac{2^{n-1}\cdot 2}{n!\,(n+1)}=\frac{2^{n-1}}{n!}\cdot\frac{2}{n+1}=x_n\cdot\frac{2}{n+1}.$$

Применяя к равенству

$$x_{n+1}=x_n\cdot\frac{2}{n+1}$$

формулу (6), получаем уравнение для a:

$$a=a\cdot\lim_{n\to\infty}\frac{2}{n+1}\Leftrightarrow a=a\cdot 0,$$

из которого следует, что предел данной последовательности равен нулю.

Пример 5. Число e как предел бесконечной числовой последовательности*). Рассмотрим последовательность, задаваемую формулой общего члена

$$x_n=\left(1+\frac{1}{n}\right)^n.$$

Докажем монотонность этой последовательности. Раскладывая выражение $\left(1+\frac{1}{n}\right)^n$ по формуле бинома Ньютона, получаем

$$x_n=\left(1+\frac{1}{n}\right)^n=1+n\cdot\frac{1}{n}+\frac{n(n-1)}{2!}\cdot\frac{1}{n^2}+\frac{n(n-1)(n-2)}{3!}\cdot\frac{1}{n^3}+\ldots$$
$$\ldots+\frac{n(n-1)(n-2)\ldots(n-n+1)}{n!}\cdot\frac{1}{n^n}=1+1+\frac{1}{2!}\left(1-\frac{1}{n}\right)+$$
$$+\frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)+\ldots+\frac{1}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{n-1}{n}\right)$$

Если теперь написать аналогичное разложение для x_{n+1}, то к разложению для x_n прежде всего добавится $(n+2)$-й положительный член; каждый же из написанных $n+1$ членов увеличится, так как любой множитель в скобках вида $1-\frac{s}{n}$ заменится бо́льшим множителем $1-\frac{s}{n+1}$. Отсюда и следует, что

$$x_{n+1} > x_n,$$

т. е. данная последовательность (x_n) возрастающая.

Докажем, что последовательность (x_n) ограничена. Опустив в выражении для x_n

$$x_n=1+1+\frac{1}{2!}\left(1-\frac{1}{n}\right)+\frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)+\ldots$$
$$\ldots+\frac{1}{n!}\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{n-1}{n}\right)$$

*) О числе e см. также п. 1.5 гл. 9.

все множители, стоящие в круглых скобках, мы тем самым увеличим правую часть равенства (так как каждый из опущенных множителей меньше единицы) и в результате получим неравенство

$$x_n < 2+\frac{1}{2!}+\frac{1}{3!}+\ldots+\frac{1}{n!}. \qquad (13)$$

Из ограниченности последовательности (12), рассмотренной в примере 4, следует, что

$$2^{n-1} \leqslant n!,$$

т. е.

$$2 \leqslant 2!, \quad 2^2 \leqslant 3!, \quad 2^3 \leqslant 4!, \ \ldots, \quad 2^{n-1} \leqslant n!;$$

и, следовательно,

$$\frac{1}{2} \geqslant \frac{1}{2!}, \quad \frac{1}{2^2} \geqslant \frac{1}{3!}, \quad \frac{1}{2^3} \geqslant \frac{1}{4!}, \ \ldots, \ \frac{1}{2^{n-1}} \geqslant \frac{1}{n!}.$$

Используя эти неравенства, получаем

$$x_n < 2+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\ldots+\frac{1}{n!} \leqslant 2+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\ldots+\frac{1}{2^{n-1}}.$$

В правой части последнего неравенства сумма всех членов, начиная со второго, есть сумма $n-1$ членов бесконечно убывающей геометрической прогрессии, которая равна 1, а потому при любом $n \in \mathbb{N}$

$$x_n < 3.$$

Таким образом, последовательность $x_n=\left(1+\frac{1}{n}\right)^n$, $n \in \mathbb{N}$, монотонно возрастающая и ограниченная и, согласно теореме Вейерштрасса, имеет предел. Этот предел имеет специальное обозначение e;

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e \approx 2{,}71828\ldots$$

2.5. Необходимое и достаточное условие сходимости последовательности. Выше были рассмотрены два способа доказательства сходимости последовательностей. Один способ был основан на использовании теорем о пределах последовательностей (см. п. 2.3), а второй — на использовании теоремы Вейерштрасса (см. п. 2.4). Однако оба эти способа доказательства сходимости последовательности применимы для исследования весьма небольшого числа довольно простых последовательностей.

Возникает вопрос: а как же доказывать сходимость, скажем, немонотонной последовательности, которая задается не аналитическим, а каким-либо другим способом (например, рекуррентным), или же последовательности, которая задается аналитической формулой, но предел которой невозможно вычислить, ограничиваясь теоремами о пределах последовательностей? Ясно, что при этом мы уже не можем пользоваться определением предела последовательности для доказательства ее сходимости, так как в этом определении уже участвует то самое число a, существование которого мы должны доказать. Кроме того, в ряде случаев нас может интересовать лишь вопрос о сходимости последовательности, но не интересовать вопрос о том, к какому числу она сходится.

Теорема, которая отвечает на вопрос о существовании предела последовательности и в формулировке которой участвует только то, что нам дано, а именно члены последовательности, была сформулирована и доказана Б. Больцано и О. Коши.

Теорема Больцано—Коши (критерий сходимости). *Для того чтобы последовательность* (x_n) *имела конечный*

предел, необходимо и достаточно, чтобы для каждого положительного числа ε *существовал такой номер* $N_0(\varepsilon)$, *чтобы неравенство*

$$|x_n - x_{n+p}| < \varepsilon$$

выполнялось для всех $n > N_0$ *и* $p \geqslant 1$ (т. е. для любых натуральных чисел n и p, удовлетворяющих этим неравенствам).

§ 3. Числовые ряды и бесконечные произведения

3.1. Понятие числового ряда. Пусть дана некоторая бесконечная числовая последовательность

$$x_1, x_2, x_3, \ldots, x_n, \ldots \tag{1}$$

Поставим перед собой задачу: вычислить сумму всех членов бесконечной последовательности (1). Вычислить сумму бесконечного числа слагаемых так, как мы вычисляем сумму конечного числа слагаемых, нельзя — мы никогда не сможем завершить процесс сложения. Однако оказывается, что, используя понятие предела последовательности, можно ввести понятие суммы бесконечного числа слагаемых. Для этого возьмем первый член последовательности (x_n) и обозначим его через S_1:

$$S_1 = x_1.$$

Сложим первый член последовательности x_1 со вторым членом x_2 и обозначим их сумму через S_2:

$$S_2 = x_1 + x_2.$$

Сложим первые три члена последовательности (1) и обозначим их сумму через S_3:

$$S_3 = x_1 + x_2 + x_3.$$

Продолжая начатую процедуру, т. е. последовательно складывая четыре, пять, шесть и т. д. первых членов последовательности (x_n), обозначим суммы четырех, пяти, шести и т. д. первых членов через $S_4, S_5, S_6, \ldots$:

$$\begin{aligned} S_4 &= x_1 + x_2 + x_3 + x_4, \\ S_5 &= x_1 + x_2 + x_3 + x_4 + x_5, \\ S_6 &= x_1 + x_2 + x_3 + x_4 + x_5 + x_6, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

Для заданной числовой последовательности (x_n) выражение

$$x_1 + x_2 + x_3 + \ldots + x_n + \ldots \tag{2}$$

называют *числовым рядом* и обозначают символом

$$\sum_{n=1}^{\infty} x_n.$$

Слагаемое x_n называется n-м *членом* ряда. Сумма n первых членов ряда называется n-й *частичной суммой* этого ряда и обозначается через S_n:

$$S_n = x_1 + x_2 + x_3 + \ldots + x_n.$$

Если существует конечный предел S последовательности (S_n) частичных сумм ряда (2), то говорят, что этот ряд *сходится*, а

указанный предел S называют *суммой* ряда (2) и пишут

$$S = x_1 + x_2 + x_3 + \ldots + x_n + \ldots = \sum_{n=1}^{\infty} x_n.$$

Если же $S_n \longrightarrow +\infty$, либо $S_n \longrightarrow -\infty$, либо предела последовательности частичных сумм (S_n) вовсе не существует, то ряд (2) называют *расходящимся*.

Остатком ряда

$$x_1 + x_2 + x_3 + \ldots + x_k + x_{k+1} + \ldots + x_{k+l} + \ldots \tag{3}$$

после k-го члена называется бесконечный ряд

$$x_{k+1} + x_{k+2} + \ldots + x_{k+l} + \ldots,$$

получающийся из ряда (3) отбрасыванием первых k членов.

Простейшим примером бесконечного числового ряда служит сумма всех членов геометрической прогрессии с первым членом, равным a, и знаменателем, равным q:

$$a + aq + aq^2 + \ldots + aq^{n-1} + \ldots; \tag{4}$$

n-я частичная сумма этого ряда есть сумма n первых членов геометрической прогрессии и вычисляется по формуле

$$S_n = \frac{a(1-q^n)}{1-q} \qquad (\text{при } q \neq 1).$$

Если знаменатель прогрессии удовлетворяет неравенству $|q| < 1$, то последовательность частичных сумм (S_n) имеет предел

$$S = \frac{a}{1-q},$$

который называют *суммой* всех членов бесконечно убывающей геометрической прогрессии (т. е. суммой ряда (4)).

Если знаменатель прогрессии удовлетворяет неравенству $|q| \geqslant 1$, то выражение (4) дает пример расходящегося ряда. При этом если $q < -1$, то предела частичных сумм не существует, если же $q \geqslant 1$, то последовательность частичных сумм стремится либо к $+\infty$, либо к $-\infty$.

Свойства бесконечных числовых рядов:

1) если числовой ряд сходится, то сходится и любой из его остатков; обратно, из сходимости остатка ряда вытекает сходимость исходного ряда;

2) если ряд сходится, то остаток ряда после k-го члена стремится к нулю при $k \longrightarrow \infty$;

3) если члены сходящегося ряда умножить на один и тот же множитель c, то его сходимость не нарушится, а сумма умножится на c:

$$\sum_{n=1}^{\infty} cx_n = c \sum_{n=1}^{\infty} x_n;$$

4) два сходящихся ряда

$$x_1 + x_2 + x_3 + \ldots + x_n + \ldots = \sum_{n=1}^{\infty} x_n = X,$$

$$y_1 + y_2 + y_3 + \ldots + y_n + \ldots = \sum_{n=1}^{\infty} y_n = Y$$

можно почленно складывать (или вычитать), и ряд

$$(x_1 \pm y_1)+(x_2 \pm y_2)+(x_3 \pm y_3)+\ldots+(x_n \pm y_n)+\ldots$$

также сходится, а его сумма (разность) равна $X+Y$ (соответственно $X-Y$);

5) общий член x_n сходящегося ряда стремится к нулю (необходимое условие сходимости).

3.2. Положительные числовые ряды. Числовой ряд

$$x_1+x_2+x_3+\ldots+x_n+\ldots=\sum_{n=1}^{\infty} x_n$$

называют *положительным*, если все члены ряда — положительные числа:

$$x_1>0,\quad x_2>0,\quad x_3>0,\quad \ldots,\quad x_n>0,\quad \ldots$$

Последовательность частичных сумм положительного ряда — возрастающая последовательность. Положительный числозой ряд сходится, если последовательность частичных сумм ряда ограничена сверху, и расходится в противном случае (необходимое и достаточное условие сходимости положительного числового ряда).

Сходимость (или расходимость) положительного числового ряда часто устанавливается путем сравнения данного числового ряда с другим, заведомо сходящимся (или расходящимся) числовым рядом. В основе этого сравнения лежит следующая теорема (свойство числовых рядов).

Пусть даны два положительных числовых ряда

$$x_1+x_2+\ldots+x_n+\ldots=\sum_{n=1}^{\infty} x_n, \tag{I}$$

$$y_1+y_2+\ldots+y_n+\ldots=\sum_{n=1}^{\infty} y_n. \tag{II}$$

Если, начиная с некоторого члена, для всех $n>N_0$ выполняется неравенство $x_n \leqslant y_n$, то из сходимости ряда (II) *следует сходимость ряда* (I), *а из расходимости ряда* (I) *следует расходимость ряда* (II).

Таким образом, для доказательства расходимости ряда

$$x_1+x_2+x_3+\ldots+x_n+\ldots=\sum_{n=1}^{\infty} x_n \tag{5}$$

необходимо построить такой заведомо расходящийся ряд

$$y_1+y_2+y_3+\ldots+y_n+\ldots=\sum_{n=1}^{\infty} y_n,$$

чтобы $y_n \leqslant x_n$ для всех $n>N_3$

Для доказательства сходимости ряда (5) необходимо построить такой заведомо сходящийся ряд

$$z_1+z_2+z_3+\ldots+z_n+\ldots=\sum_{n=1}^{\infty} z_n,$$

чтобы $z_n \geqslant x_n$ для всех $n>N_4$.

В общем случае трудно дать какие-либо конкретные рекомендации, как строить указанные ряды $\sum_{n=1}^{\infty} y_n$ и $\sum_{n=1}^{\infty} z_n$. Часто в качестве таких рядов удобно брать ряды, образованные из членов геометрической прогрессии, которые в зависимости от значения знаменателя могут быть как сходящимися, так и расходящимися, или гармонические ряды.

Числовой ряд

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots+\frac{1}{n}+\ldots=\sum_{n=1}^{\infty}\frac{1}{n}$$

называется *гармоническим рядом*. Каждый член такого ряда, начиная со второго, представляет собой среднее гармоническое двух соседних членов (число c, удовлетворяющее равенству $\frac{1}{c}=\frac{1}{2}\left(\frac{1}{a}+\frac{1}{b}\right)$, называется *средним гармоническим* чисел a и b).

Числовой ряд более общего вида

$$1+\frac{1}{2^s}+\frac{1}{3^s}+\frac{1}{4^s}+\ldots+\frac{1}{n^s}+\ldots=\sum_{n=1}^{\infty}\frac{1}{n^s}, \tag{6}$$

где s — любое действительное число, часто также называют гармоническим рядом

Гармонический ряд вида (6) расходится при $s \leqslant 1$ и сходится при $s > 1$.

Сходимость (или расходимость) рядов может быть также установлена с помощью следующих следствий сформулированной выше теоремы.

Пусть

$$x_1+x_2+\ldots+x_n+\ldots, \tag{I}$$
$$y_1+y_2+\ldots+y_n+\ldots \tag{II}$$

— два положительных ряда.

Тогда:

1) если существует предел

$$\lim_{n\to\infty}\frac{x_n}{y_n}=A, \quad A \geqslant 0,$$

то из сходимости ряда (II) при конечном A вытекает сходимость ряда (I), а из расходимости ряда (I) при $A > 0$ вытекает расходимость ряда (II), т. е. при конечном значении A оба ряда сходятся и расходятся одновременно;

2) если, начиная с какого-то номера N_0, для всех $n > N_0$ выполняется неравенство

$$\frac{x_{n+1}}{x_n} \leqslant \frac{y_{n+1}}{y_n},$$

то из сходимости ряда (II) вытекает сходимость ряда (I), а из расходимости ряда (I) вытекает расходимость ряда (II).

3.3. Признаки сходимости числовых рядов. Сходимость (или расходимость) ряда

$$x_1+x_2+\ldots+x_n+\ldots$$

может быть установлена с помощью следующих признаков.

Признак Коши. Если предел

$$\lim_{n\to\infty}\sqrt[n]{x_n}$$

существует и равен K, то при $K<1$ ряд сходится, а при $K>1$ расходится. Если $K=1$, то ряд может быть как сходящимся, так и расходящимся.

Примеры. 1. Ряд $\sum\limits_{n=1}^{\infty}\frac{1}{2^n}$ сходящийся, так как по признаку Коши

$$\lim_{n\to\infty}\sqrt[n]{\frac{1}{2^n}}=\frac{1}{2}<1.$$

2. Ряд $\sum\limits_{n=1}^{\infty}\left(\frac{2n}{n+1}\right)^n$ расходящийся, так как по признаку Коши

$$\lim_{n\to\infty}\sqrt[n]{\left(\frac{2n}{n+1}\right)^n}=\lim_{n\to\infty}\frac{2n}{n+1}=2>1.$$

Признак Даламбера. Если предел

$$\lim_{n\to\infty}\frac{x_{n+1}}{x_n}$$

существует и равен D, то при $D<1$ ряд сходится, а при $D>1$ расходится. При $D=1$ ряд может быть как сходящимся, так и расходящимся.

Примеры. 1. Ряд $\sum\limits_{n=1}^{\infty}\frac{2^n}{n!}$ сходится, так как по признаку Даламбера

$$\lim_{n\to\infty}\frac{x_{n+1}}{x_n}=\lim_{n\to\infty}\frac{2^{n+1}}{(n+1)!}\cdot\frac{n!}{2^n}=\lim_{n\to\infty}\frac{2}{n+1}=0<1.$$

2. Ряд $\sum\limits_{n=1}^{\infty}n!\left(\frac{10}{n}\right)^n$ расходится, так как

$$\lim_{n\to\infty}\frac{x_{n+1}}{x_n}=\lim_{n\to\infty}\frac{(n+1)!\,10^{n+1}}{(n+1)^{n+1}}\cdot\frac{n^n}{n!\,10^n}=$$

$$=\lim_{n\to\infty}\frac{(n+1)\cdot 10\cdot n^n}{(n+1)^{n+1}}=\lim_{n\to\infty}\frac{10}{\left(1+\frac{1}{n}\right)^n}=\frac{10}{e}>1.$$

Признак Раабе. Если предел

$$\lim_{n\to\infty}\left[n\left(1-\frac{x_{n+1}}{x_n}\right)\right]$$

существует и равен R, то при $R>1$ ряд сходится, а при $R<1$ расходится. При $R=1$ ряд может быть как сходящимся, так и расходящимся.

Признак Раабе значительно «сильнее» признака Даламбера — если признак Даламбера дает ответ на вопрос о сходимости ряда, то и признак Раабе его также дает. Если $D<1$, то $R=+\infty$, если $D>1$, то $R=-\infty$, т. е. все возможные значения D описываются всего двумя значениями $R=\pm\infty$. Все остальные значения R (кроме $R=1$), также дающие ответ на вопрос о сходимости ряда, соответствуют случаям, когда $D=1$, т. е. случаям, когда признак Даламбера не дает ответа на вопрос о сходимости ряда.

Так, например, признак Даламбера для ряда $\sum_{n=1}^{\infty}\frac{1}{n^2}$ не дает ответа на вопрос о его сходимости, в то время как по признаку Раабе

$$\lim_{n\to\infty}\left[n\left(1-\frac{x_{n+1}}{x_n}\right)\right]=\lim_{n\to\infty}\left[n\left(1-\frac{n^2}{(n+1)^2}\right)\right]=\lim_{n\to\infty}\left[n\cdot\frac{2n+1}{n^2+2n+1}\right]=$$

$$=\lim_{n\to\infty}\frac{2n^2+n}{n^2+2n+1}=\lim_{n\to\infty}\frac{2+\frac{1}{n}}{1+\frac{2}{n}+\frac{1}{n^2}}=2,$$

т. е. данный ряд сходится.

3.4. Некоторые конечные суммы и числовые ряды.

$$1+2+3+\ldots+n=\frac{n(n+1)}{2};$$

$$1+3+5+\ldots+(2n-1)=n^2;$$

$$2+4+6+\ldots+2n=n(n+1);$$

$$1^2+2^2+3^2+\ldots+n^2=\frac{n(n+1)(2n+1)}{6};$$

$$1^2+3^2+5^2+\ldots+(2n-1)^2=\frac{n(4n^2-1)}{3};$$

$$\sum_{n=0}^{\infty}\frac{1}{2^n}=1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots=2;$$

$$\sum_{n=0}^{\infty}(-1)^n\frac{1}{2^n}=1-\frac{1}{2}+\frac{1}{4}-\frac{1}{8}+\ldots=\frac{2}{3};$$

$$\sum_{n=0}^{\infty}(-1)^n\frac{1}{2n+1}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\frac{1}{9}-\ldots=\frac{\pi}{4};$$

$$\sum_{n=1}^{\infty}\frac{1}{n(n+1)}=\frac{1}{1\cdot2}+\frac{1}{2\cdot3}+\frac{1}{3\cdot4}+\ldots=1;$$

$$\sum_{n=1}^{\infty}\frac{1}{(2n-1)(2n+1)}=\frac{1}{1\cdot3}+\frac{1}{3\cdot5}+\frac{1}{5\cdot7}+\ldots=\frac{1}{2};$$

$$\sum_{n=2}^{\infty}\frac{1}{(n-1)(n+1)}=\frac{1}{1\cdot3}+\frac{1}{2\cdot4}+\frac{1}{3\cdot5}+\ldots=\frac{3}{4};$$

$$\sum_{n=1}^{\infty} \frac{1}{(4n-1)(4n+1)} = \frac{1}{3\cdot 5} + \frac{1}{7\cdot 9} + \frac{1}{11\cdot 13} + \ldots = \frac{1}{2} - \frac{\pi}{8};$$

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+2)} = \frac{1}{1\cdot 2\cdot 3} + \frac{1}{2\cdot 3\cdot 4} + \ldots = \frac{1}{4};$$

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \ldots = \frac{\pi^2}{6};$$

$$\sum_{n=1}^{\infty} (-1)^n \frac{1}{n^2} = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \ldots = \frac{\pi^2}{12};$$

$$\sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots = e;$$

$$\sum_{n=0}^{\infty} (-1)^n \frac{1}{n!} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots = \frac{1}{e};$$

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \ln 2;$$

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)!} = \sin 1;$$

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!} = \cos 1 - 1.$$

3.5. Понятие бесконечного произведения. Рассмотрим бесконечную числовую последовательность $x_1, x_2, x_3, \ldots, x_n, \ldots$ Образуем из данной числовой последовательности последовательность, членами которой являются числа

$$X_1 = x_1, \quad X_2 = x_1 \cdot x_2, \quad X_3 = x_1 \cdot x_2 \cdot x_3, \ \ldots,$$
$$X_n = x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n, \ \ldots$$

Для заданной числовой последовательности (x_n) выражение

$$x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n \cdot \ldots \tag{7}$$

называется *бесконечным произведением*. Бесконечное числовое произведение (7) обозначается

$$\prod_{n=1}^{\infty} x_n.$$

Произведение n первых членов последовательности (x_n) называется *частичным произведением* данного бесконечного произведения:

$$X_n = x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n.$$

Если существует конечный и отличный от нуля предел X последовательности (X_n) частичных произведений, то говорят, что бес-

конечное произведение *сходится*; предел X называют *значением* бесконечного произведения (1) и пишут

$$X = x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n \cdot \ldots$$

В противном случае (когда конечный предел последовательности частичных произведений не существует или равен нулю) бесконечное произведение называют *расходящимся*.

3.6. **Связь бесконечных произведений с рядами.** Рассмотрим бесконечную числовую последовательность (x_n), все члены которой положительны. Составим для этой последовательности бесконечное произведение

$$x_1 \cdot x_2 \cdot x_3 \cdot \ldots \cdot x_n \cdot \ldots = \prod_{n=1}^{\infty} x_n. \tag{8}$$

Для сходимости бесконечного произведения (8) необходимо и достаточно, чтобы сходился бесконечный ряд

$$\sum_{n=1}^{\infty} \ln x_n. \tag{9}$$

Если ряд (9) сходится к числу S, то бесконечное произведение (8) сходится к числу $X = e^S$.

§ 4. Прогрессии

4.1. **Арифметическая прогрессия.** Последовательность, у которой задан первый член a_1, а каждый следующий член, начиная со второго, равен предшествующему члену, сложенному с одним и тем же числом d, называется *арифметической прогрессией*. Число d называется *разностью* арифметической прогрессии. Если число $d > 0$, то прогрессия называется *возрастающей*, если $d < 0$ — *убывающей*.

Для того чтобы последовательность (a_n) была арифметической прогрессией, необходимо и достаточно, чтобы для любого $n > 1$ выполнялось равенство

$$a_n = \frac{a_{n-1} + a_{n+1}}{2}.$$

Если заданы первый член a_1 и разность арифметической прогрессии d, то n-й член арифметической прогрессии вычисляется по формуле

$$a_n = a_1 + d\,(n-1).$$

Эта формула называется *формулой общего члена* арифметической прогрессии. Члены арифметической прогрессии a_{n-l}, a_n и a_{n+l} $(n > l)$ связаны между собой равенством

$$a_n = \frac{a_{n+l} + a_{n-l}}{2}.$$

Сумма первых n членов арифметической прогрессии (a_n) вычисляется по формуле

$$S_n = \frac{a_1 + a_n}{2} \cdot n$$

или по формуле

$$S_n = \frac{2a_1 + d(n-1)}{2} \cdot n.$$

Ряд, образованный из членов арифметической прогрессии

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots = \sum_{n=1}^{\infty} a_n,$$

при любых ненулевых числовых значениях первого члена a_1 и разности d будет расходящимся.

4.2. Геометрическая прогрессия. Последовательность, у которой задан первый член $u_1 \neq 0$, а каждый следующий член, начиная со второго, получается умножением предыдущего члена на одно и то же постоянное для данной последовательности отличное от нуля число q, называется *геометрической прогрессией*. Число q называется знаменателем геометрической прогрессии.

Для того чтобы последовательность (u_n) была геометрической прогрессией, необходимо и достаточно, чтобы для любого $n > 1$ выполнялось равенство

$$u_n^2 = u_{n-1} \cdot u_{n+1}.$$

Если заданы первый член u_1 и знаменатель геометрической прогрессии q, то n-й член геометрической прогрессии вычисляется по формуле

$$u_n = u_1 \cdot q^{n-1}.$$

Эта формула называется *формулой общего члена* геометрической прогрессии. Члены геометрической прогрессии u_{n-l}, u_n и u_{n+l} $(n > l)$ связаны между собой равенством

$$u_n^2 = u_{n-l} \cdot u_{n+l}.$$

Сумма первых n членов геометрической прогрессии вычисляется по формуле

$$S_n = \frac{u_1(1-q^n)}{1-q}.$$

Бесконечный числовой ряд, образованный из членов геометрической прогрессии

$$u_1 + u_2 + u_3 + \ldots + u_n + \ldots = \sum_{n=1}^{\infty} u_n, \qquad (1)$$

при $|q| < 1$ сходится и его сумма S равна

$$S = \frac{u_1}{1-q}, \qquad (2)$$

при $|q| \geqslant 1$ ряд (1) расходится.

Формулу (2) также называют формулой *суммы членов бесконечно убывающей геометрической прогрессии*. Здесь под «бесконечно убывающей геометрической прогрессией» понимается прогрессия, члены которой убывают по абсолютной величине, т. е.

$$|u_1| > |u_2| > |u_3| > \ldots > |u_n| > \ldots$$

§ 5. Числовые функции

5.1. Понятие числовой функции. Пусть X и Y — два множества действительных чисел. Элементы этих множеств будем обозначать буквами x и y соответственно. Задать *соответствие* между множеством X и множеством Y — это значит указать правило, по которому для каждого числа x из множества X выбирается одно, несколько или даже бесконечно много чисел y из множества Y. При этом может оказаться, что некоторым числам $x \in X$ не будет соответствовать ни одно из чисел множества Y.

Например, если задать соответствие между множествами действительных чисел $X = \mathbf{R}$ и $Y = \mathbf{R}$ формулой

$$x^2 + y^2 = 1,$$

то каждому числу x из промежутка $(-1; 1)$ будут соответствовать два числа:

$$y = \sqrt{1 - x^2} \quad \text{и} \quad y = -\sqrt{1 - x^2},$$

Числам $x = 1$ и $x = -1$ соответствует единственное число $y = 0$, а всем остальным $x \in \mathbf{R}$ не соответствует никакое y.

Соответствия между числовым множеством X и числовым множеством Y обычно обозначают буквами латинского алфавита f, F, g, ... и пишут

$$X \xrightarrow{f} Y, \quad X \xrightarrow{F} Y, \quad X \xrightarrow{g} Y, \ldots,$$
$$f\colon X \longrightarrow Y, \quad F\colon X \longrightarrow Y, \quad g\colon X \longrightarrow Y, \ldots,$$

Соответствие f, относящее каждому данному числу x из множества X единственное число y из множества Y, называют *числовой функцией* аргумента x и пишут

$$y = f(x). \tag{1}$$

При этом x называют *независимой переменной*, а y — *зависимой переменной*. Буква f в записи (1) характеризует именно то правило, по которому получается значение y, отвечающее заданному значению x. Множество X называют *областью определения* функции $f(x)$, а множество Y_1 $(Y_1 \subset Y)$, образованное из всевозможных чисел y, вычисляемых по формуле (1), — *областью изменения* (или *множеством значений*) функции $f(x)$. Иногда область определения и область изменения функции $y = f(x)$ обозначают $D(f)$ и $E(f)$ соответственно.

В определении понятия функции существенно, что, задавая правило, по которому для заданного значения x получается значение y, необходимо также указать область определения функции, т. е. множество X. (Множество Y_1 — область изменения функции — обычно не указывается, поскольку сам закон соответствия уже определяет множество значений функции.) В тех случаях, когда область определения функции явно не указана, под областью определения функции понимается такое множество X, что для каждого $x \in X$ определено правило выбора единственного значения y.

Так, например, пусть функция задается равенством

$$y = \sqrt{x}.$$

Поскольку в множестве действительных чисел квадратные корни можно извлекать лишь из неотрицательных чисел, то область определения этой функции — множество всех чисел, принадлежащих промежутку $[0; +\infty)$.

Или, например, пусть функция задается равенством

$$y=\frac{1}{x}.$$

У этой функции область определения — любое действительное число, отличное от нуля, поскольку операция деления на нуль не определена.

Числовая функция $y=f(x)$ может быть определена и как множество упорядоченных пар действительных чисел $(x;\ f(x))$ таких, что для каждого x в множестве этих пар имеется не более одной пары с данным первым элементом x. Например, функция, задаваемая формулой $y=x^2$, может быть также задана как множество всех упорядоченных пар вида $(x;\ x^2)$.

Если переменные x и y рассматривать как декартовы координаты точек на плоскости, то *графиком функции* $y=f(x)$ называют множество точек координатной плоскости Oxy с координатами $(x;\ f(x))$.

5.2. Способы задания функций. В математическом анализе функции обычно задаются так называемым аналитическим способом.

Основные формы аналитического задания функции.

1. Явная форма задания функции. Функция задается в виде формулы, указывающей операции (и последовательность их выполнения), которые необходимо совершить над значением независимой переменной, в результате которых получается значение зависимой переменной.

Например, пусть функция $y=f(x)$ имеет вид

$$y=(\sqrt{x}-1)^2.$$

Область определения данной функции $D(f)=[0;\ +\infty)$.

Возьмем некоторое значение $x_0\in D(f)$. Для получения значения $y_0=f(x_0)$ над величиной x_0 следует произвести следующие операции, в результате которых получится величина y_0: а) извлечь квадратный корень из числа x_0; б) вычесть из полученного значения квадратного корня единицу; в) полученную разность возвести в квадрат.

При аналитическом способе задания функции при одних значениях аргумента $x\in D_1$ функция может задаваться одной формулой, а при других значениях аргумента $x\in D_2$ — другой формулой ($D_1\cup D_2=D(f)$ и $D_1\cap D_2=\varnothing$).

Примером такой функции, определенной на всей числовой оси, является функция, определяемая следующими условиями:

$$\begin{aligned} &f(x)=1, && \text{если} \quad x>0,\\ &f(x)=-1, && \text{если} \quad x<0,\\ &f(0)=0. \end{aligned}$$

Эта функция имеет специальное обозначение:

$$\operatorname{sign} x,$$

читается «сигнум x» и называется «знак числа x».

2. Неявная форма задания функции. Под неявным заданием функции понимается задание функции в виде уравнения:

$$F(x,\ y)=0. \tag{2}$$

Уравнение (2) задает функцию лишь в том случае, если множество упорядоченных пар $(x;\ y)$, являющихся решением данного уравнения, таково, что для любого числа x имеется в этом множестве не более

одной пары $(x_0; y_0)$ с первым элементом x_0. Например, уравнение

$$xy - 1 = 0$$

задает функцию, в то время как уравнение вида

$$x^2 + y^2 = 1$$

задает не функцию, а соответствие, так как среди множества упорядоченных пар $(x; y)$, являющихся решением данного уравнения, найдутся, например, две такие пары с совпадающим первым элементом:

$$\left(\frac{\sqrt{3}}{2}; \frac{1}{2}\right) \text{ и } \left(\frac{\sqrt{3}}{2}; -\frac{1}{2}\right).$$

3. Параметрическая форма задания функции. При параметрическом задании функции соответствующие друг другу значения переменных x и y выражаются через третью величину, называемую *параметром*:

$$x = \varphi(t), \quad y = \psi(t).$$

В некоторых случаях функция, заданная в параметрическом виде, может быть также записана и в явной форме. Например, функция, заданная в параметрической форме:

$$x = \sin t, \quad y = \cos t, \quad t \in [-\pi/2, \pi/2],$$

допускает запись в явной форме:

$$y = \sqrt{1 - x^2}.$$

Кроме аналитического способа задания, функция может быть также задана:

1) таблицей своих значений или правилом вычисления таблицы;
2) графиком;
3) перечислением характерных свойств функции и т. д.

5.3. Сумма, произведение, разность и частное двух функций. Функции $f(x)$ и $g(x)$ называются *равными* на общей части их области определения $D(f) \cap D(g)$, если $f(x) = g(x)$ при всех $x \in D(f) \cap D(g)$.

Суммой двух функций $f(x)$ и $g(x)$, определенных на множествах значений независимой переменной $D(f)$ и $D(g)$ соответственно, называется функция $S(x)$, задаваемая условиями:

1) область определения функции $S(x)$ (если нет специальной оговорки) есть общая часть множества $D(f)$ и $D(g)$, т. е.

$$D(S) = D(f) \cap D(g);$$

2) значение функции $S(x)$ в каждой точке $x_0 \in D(S)$ вычисляется как сумма значений функций $f(x)$ и $g(x)$ в точке x_0:

$$S(x_0) = f(x_0) + g(x_0).$$

Аналогично определяются разность, произведение и частное двух функций, причем частное двух функций

$$\frac{f(x)}{g(x)}$$

определено лишь при тех значениях $x \in D(f) \cap D(g)$, при которых функция $g(x)$ не обращается в нуль.

5.4. Сложная функция (суперпозиция функций). Пусть $y = f(x)$ — числовая функция с областью определения $D(f)$ и областью измене-

ния (множеством значений) $E(f)$, а $z=g(y)$ — числовая функция, заданная на множестве $E(f)$ или некотором его подмножестве, с областью изменения $E(g)$. Соответствие, относящее каждому данному числу x из множества $D(f)$ единственное число y из множества $E(f)$, а числу y — единственное число z из множества $E(g)$, называют *сложной функцией* (или *суперпозицией* функций $y=f(x)$ и $z=g(y)$) и пишут

$$z=g(f(x)).$$

Большинство функций, изучаемых в элементарной математике и математическом анализе, можно рассматривать как сложные функции; например, функция

$$z=\sqrt{x}-1$$

может быть представлена как суперпозиция двух функций:

$$y=\sqrt{x}, \quad z=y-1.$$

5.5. Четные и нечетные функции. Числовая функция $y=f(x)$ называется *четной*, если:

1) область определения функции симметрична относительно точки O числовой оси (т. е. если точка x_0 принадлежит области определения функции, то и точка $-x_0$ также принадлежит области определения функции);

2) для любого значения независимой переменной, принадлежащего области определения функции, выполняется равенство

$$f(x)=f(-x). \tag{3}$$

Числовая функция $y=f(x)$ называется *нечетной*, если:

1) область определения функции симметрична относительно точки O числовой оси (т. е. если точка x_0 принадлежит области определения функции, то и точка $-x_0$ также принадлежит области определения функции);

2) для любого значения независимой переменной, принадлежащего области определения функции, выполняется равенство

$$f(x)=-f(-x). \tag{3'}$$

График четной функции симметричен относительно оси ординат, а график нечетной функции симметричен относительно начала координат, т. е. центрально симметричен.

Свойства четных и нечетных функций:

1) сумма, разность, произведение и частное двух четных функций — четная функция;

2) сумма и разность нечетных функций — нечетная функция, а произведение и частное — четная функция.

Доказательство четности (или нечетности) функции $y=f(x)$ проводится следующим образом. Выясняется, симметрична ли область определения функции $y=f(x)$ относительно точки O числовой оси. Если область определения функции не симметрична относительно точки O, то функция не является ни четной, ни нечетной. Если же область определения функции симметрична относительно точки O, то переходят к проверке справедливости равенств (3) (или (3')).

Если выполняется равенство (3), то функция $f(x)$ четная; если (3'), то нечетная. Если не выполняется ни одно из приведенных равенств, то функция не является ни четной, ни нечетной.

Примеры. 1. Функция $f(x)=\frac{1}{x+1}$ не является четной и не является нечетной, так как ее область определения не симметрична относительно точки O (в точке $x=1$ функция определена, а в точке $x=-1$ не определена).

2. Область определения функции $f(x)=\frac{x^2+x}{x}$ симметрична относительно точки O, но функция не является ни четной, ни нечетной, так как

$$f(-x)=\frac{(-x)^2+(-x)}{-x}=\frac{x^2-x}{-x}=-\frac{x^2-x}{x},$$

и, следовательно,

$$f(-x)\neq f(x) \quad \text{и} \quad f(x)\neq -f(-x).$$

3. Область определения функции $f(x)=\frac{1}{x^2-1}$ симметрична относительно точки O, и, как нетрудно проверить, функция удовлетворяет равенству (3):

$$f(-x)=\frac{1}{(-x)^2-1}=\frac{1}{x^2-1}=f(x).$$

Следовательно, функция $f(x)=\frac{1}{x^2-1}$ четная.

5.6. Периодические функции. Функция $f(x)$ называется *периодической* с периодом T (T — некоторое действительное число, отличное от нуля), если:

1) для любого значения аргумента x из области определения функции значения аргумента $x+T$ и $x-T$ также принадлежат области определения функции;

2) при любом x из области определения функции

$$f(x\pm T)=f(x).$$

Обычно под *периодом* функции понимают наименьший из всех положительных периодов (если такой существует). В этом случае все периоды функции кратны ее наименьшему периоду:

$$T=kT_0,$$

где T_0 — наименьший положительный период функции, а k — целое число, отличное от нуля.

Существуют периодические функции, не имеющие наименьшего положительного периода. Примером такой функции является функция Дирихле, обозначаемая $D(x)$ и задаваемая условиями:

$$D(x)=\begin{cases}0, & \text{если } x \text{ рационально,}\\ 1, & \text{если } x \text{ иррационально.}\end{cases}$$

Доказательство периодичности (или непериодичности) функции заключается в нахождении ее периода (или в доказательстве того, что не существует числа T, удовлетворяющего условиям 1), 2)).

Пример 1. Найти наименьший период функции $f(x)=\sin(ax+b)$, где $a\neq 0$.

Предположим, что данная функция периодическая, т. е. существует число T такое, что выполняется равенство

$$\sin(ax+b)=\sin[a(x+T)+b].$$

Проверим справедливость нашего предположения.

Решая уравнение $\sin(ax+b)=\sin[a(x+T)+b]$ относительно величины T, получим

$$T=\frac{\pi+2\pi k}{a}-2x-2\cdot\frac{b}{a}, \quad k\in Z, \tag{4}$$

$$T=\frac{2\pi n}{a}, \quad n\in Z. \tag{5}$$

Величина T, определяемая формулой (4), не удовлетворяет определению периода, так как T зависит от x. Формула (5) задает бесконечное множество чисел $T=\frac{2\pi n}{a}$ $(n\in Z)$. Наименьшим положительным числом из них является число

$$T=\frac{2\pi}{|a|}.$$

По определению функции синус, при любом x значения $x+\frac{2\pi}{|a|}$ и $x-\frac{2\pi}{|a|}$ также принадлежат области определения функции, т. е. условие 1) выполнено. Условие 2) выполнено в силу способа нахождения числа T.

Таким образом, доказано, что число $T=\frac{2\pi}{|a|}$ удовлетворяет определению периода, и наше предположение, что функция $\sin(ax+b)$ периодическая, справедливо.

Пример 2. Доказать непериодичность функции $f(x)=\sin(x^2)$.

Предположим, что существует число $T\neq 0$ такое, что $\sin(x^2)=\sin(x+T)^2$ при любом x. Однако, решая это уравнение относительно величины T, нетрудно проверить, что T зависит от x. Это означает, что не существует положительного числа T (одного и того же для всех x), для которого выполнялось бы равенство

$$\sin(x^2)=\sin(x+T)^2;$$

следовательно, данная функция непериодическая.

5.7. Ограниченные функции. Функция $f(x)$ называется *ограниченной сверху*, если существует такое число M, что для всех значений аргумента из области определения функции выполняется неравенство $f(x)\leqslant M$, и *ограниченной снизу*, если существует такое число m, что для всех значений аргумента из области определения функции выполняется неравенство $f(x)\geqslant m$.

Функция, ограниченная сверху и снизу, называется *ограниченной*.

Примеры. 1. Функция x^2 ограничена снизу, например числом 0, и не ограничена сверху.

2. Функция $-\sqrt{x}$ ограничена сверху, например числом 1, и не ограничена снизу.

3. Функция $\sin x$ ограничена сверху, например числом 1, и снизу, например числом -1.

Примеры неограниченных функций:

$$f(x)=x^3, \quad f(x)=\lg x, \quad f(x)=\frac{1}{x}.$$

5.8. Монотонные функции. Пусть $f(x)$ определена в некотором промежутке $(a;\ b)$.

Функция $f(x)$ называется *возрастающей* в промежутке $(a;\ b)$, если для любой пары значений $x_1\in(a;\ b)$, $x_2\in(a;\ b)$ из неравенства $x_1>x_2$ следует неравенство $f(x_1)>f(x_2)$, т. е. большему значению аргумента соответствует большее значение функции. Если из неравенства $x_1>x_2$ следует неравенство $f(x_1)\geqslant f(x_2)$, то функция называется *неубывающей*.

Функция $f(x)$ называется *убывающей* в промежутке $(a; b)$, если для любой пары значений $x_1 \in (a; b)$, $x_2 \in (a; b)$ из неравенства $x_1 > x_2$ следует неравенство $f(x_1) < f(x_2)$, т. е. большему значению аргумента соответствует меньшее значение функции. Если из неравенства $x_1 > x_2$ следует неравенство $f(x_1) \leqslant f(x_2)$, то функция называется *невозрастающей*. Возрастающие и убывающие функции называются *строго монотонными* функциями. Возрастающие, невозрастающие, убывающие и неубывающие функции называются *монотонными*.

Пусть функция $f(x)$ не монотонна во всей своей области определения, но область определения может быть разбита на некоторое (может быть, и бесконечное) число непересекающихся промежутков, таких, что на каждом из этих промежутков функция монотонна (на одних, возможно,— возрастающая, на других — убывающая). Такие промежутки называют *промежутками монотонности* функции.

Примеры. 1. Функция $y = x^2$ не монотонна на всей числовой оси, однако числовая ось может быть разбита на два промежутка монотонности функции:

$$(-\infty; 0) \quad \text{и} \quad [0; +\infty).$$

2. Функция $y = \operatorname{tg} x$, определенная на всей числовой оси, за исключением точек $x = \frac{\pi}{2} + \pi k$ $(k \in Z)$, не монотонна. Область определения функции может быть разбита на бесконечное число промежутков $\left(-\frac{\pi}{2} + \pi k; \frac{\pi}{2} + \pi k\right)$, на каждом из которых функция $y = \operatorname{tg} x$ возрастающая.

Доказательство монотонности функции состоит в проверке справедливости какого-либо из неравенств

$$f(x_1) > f(x_2), \quad f(x_1) \geqslant f(x_2), \quad f(x_1) < f(x_2), \quad f(x_1) \leqslant f(x_2).$$

3. Доказать монотонность функции $f(x) = \sqrt{x}$. Область определения функции — промежуток $[0; +\infty)$.

Рассмотрим разность

$$f(x_1) - f(x_2) = \sqrt{x_1} - \sqrt{x_2} \tag{6}$$

где x_1 и x_2 — значения аргумента из области определения функции и $x_1 > x_2$. Разность $\sqrt{x_1} - \sqrt{x_2}$ можно представить в виде

$$\sqrt{x_1} - \sqrt{x_2} = \frac{(\sqrt{x_1} - \sqrt{x_2})(\sqrt{x_1} + \sqrt{x_2})}{\sqrt{x_1} + \sqrt{x_2}} = \frac{x_1 - x_2}{\sqrt{x_1} + \sqrt{x_2}}.$$

Так как $x_1 > x_2$, то $\sqrt{x_1} + \sqrt{x_2} > 0$ и $\sqrt{x_1} - \sqrt{x_2} > 0$. Из последнего неравенства и равенства (6) при $x_1 > x_2$ следует неравенство $f(x_1) > f(x_2)$, и, следовательно, функция $f(x) = \sqrt{x}$ возрастающая.

Достаточное условие монотонности функции. Пусть функция $f(x)$ определена и дифференцируема на промежутке $(a; b)$. Для возрастания функции $f(x)$ на промежутке $(a; b)$ достаточно, чтобы производная данной функции была положительна при всех $x \in (a; b)$, а для убывания функции достаточно, чтобы производная была отрицательна при всех $x \in (a; b)$.

Сформулированное условие монотонности не является необходимым — можно привести примеры монотонных функций, производная которых обращается в нуль в конечном (и даже бесконечном) числе точек промежутка $(a; b)$. Простейшим примером такой функции будет функция

$$f(x) = x^3.$$

Она возрастает при всех $x \in \mathbf{R}$ и имеет производную, равную нулю, в точке 0.

5.9. Понятие обратной функции. Пусть $y = f(x)$ — числовая функция с областью определения $D(f)$ и областью изменения (множеством значений) $E(f)$. Далее для простоты будем полагать, что множества $D(f)$ и $E(f)$ представляют собой некоторые промежутки.

Выберем какое-нибудь значение y_0 из множества значений функции $E(f)$; тогда в множестве $D(f)$ обязательно найдется хотя бы одно значение $x = x_0$ такое, что

$$f(x_0) = y_0.$$

Если функция $y = f(x)$ непрерывная и строго монотонная, то для каждого значения y_0 из множества значений функции $E(f)$ найдется единственное значение x_0 из области определения функции $D(f)$ такое, что

$$f(x_0) = y_0.$$

Соответствие, относящее каждому данному числу y_0 из множества $E(f)$ единственное число x_0 из множества $D(f)$, называют функцией, *обратной* функции f, обозначают символом f^{-1} и пишут

$$x = f^{-1}(y).$$

В последней записи y — независимая переменная, а x — зависимая.

Функция f^{-1}, обратная непрерывной и строго монотонной функции f, также непрерывна и строго монотонна. Функции f и f^{-1} называют *взаимно обратными*.

Если функция $x = f^{-1}(y)$ является обратной для функции $y = f(x)$, то очевидно, что графики этих функций совпадают. Можно потребовать, чтобы и аргумент обратной функции обозначался буквой x, а зависимая переменная — буквой y, т. е. рассматривать функцию $y = f^{-1}(x)$. Однако такое переобозначение не изменит график функции, так как при этом мы обязаны переобозначить и координатные оси, т. е. вертикальную ось называть осью Ox, а горизонтальную — осью Oy. Если же дополнительно потребовать, чтобы, как обычно, новая ось Ox была горизонтальной, а новая ось Oy — вертикальной, то мы должны поменять эти оси местами, и эта замена уже изменит график функции. Проще всего указанную замену осей произвести с помощью осевой симметрии координатной плоскости относительно биссектрисы первого и третьего координатных углов. Как следствие этого преобразования получаем, что если изображать графики двух взаимно обратных функций $y = f(x)$ и $y = f^{-1}(x)$ в одной и той же системе координат Oxy, то эти графики будут симметричны относительно биссектрисы первого и третьего координатных углов.

Может оказаться, что на некотором промежутке $(a; b)$ функция непрерывна, но не монотонна. Если промежуток $(a; b)$ можно разбить на n непересекающихся промежутков

$$X_1,\ X_2,\ X_3,\ \ldots,\ X_n,\quad X_1 \cup X_2 \cup \ldots \cup X_n = (a; b)$$

таким образом, чтобы на каждом из этих промежутков данная функция была монотонна, то на каждом из промежутков X_1, X_2, X_3, ..., X_n данная функция будет иметь обратную функцию. Например, функция $y = f(x) = x^2$ с областью определения $(-\infty; +\infty)$ не монотонна и поэтому не имеет обратной функции. Если разбить область определения на два промежутка $(-\infty; 0)$ и $[0; +\infty)$, то на каждом из этих промежутков функция будет непрерывна и монотонна, а потому на каждом из указанных промежутков для данной функции

будет существовать обратная функция. На промежутке $(-\infty, 0)$ такой функцией будет функция $x = -\sqrt{y}$, а на промежутке $[0; +\infty)$ — функция $x = \sqrt{y}$. (Здесь y — независимая переменная, а x — зависимая.) Графики прямой и обратной функций на промежутке $[0; +\infty)$ после указанных выше переобозначений переменных изображены на рис. 8.4.

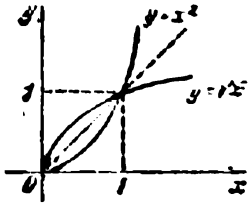


Рис. 8.4

Свойства взаимно обратных функций:

1) если функция f^{-1} является обратной для функции f, то и функция f будет обратной для функции f^{-1};

2) область определения функции f является областью изменения функции f^{-1}, а область изменения функции f — областью определения функции f^{-1};

3) графики взаимно обратных функций симметричны относительно биссектрисы первого и третьего координатных углов координатной плоскости Oxy.

5.10. Понятие числовой функции нескольких независимых переменных. В различных разделах математики и физики встречаются соотношения, в которых участвуют несколько переменных величин. Так, например, из курса геометрии известно, что объем цилиндра V вычисляется по формуле

$$V = \pi R^2 H,$$

где R — радиус основания цилиндра, H — его высота. Эта формула дает простейший пример числовой функции двух переменных — объем зависит как от радиуса основания цилиндра, так и от его высоты.

Из курса физики известна формула Клапейрона

$$pV = RT,$$

связывающая три физические характеристики состояния газа: давление p, удельный объем V и абсолютную температуру T (R — универсальная газовая постоянная). Задавая две из трех указанных величин, можно получить вполне определенное значение третьей величины. Таким образом, при использовании этой формулы мы опять встречаемся с понятием функции нескольких независимых переменных.

Переменная z называется числовой *функцией двух независимых переменных* (аргументов) x и y, заданной на множестве упорядоченных пар чисел M, если каждой упорядоченной паре чисел $(x; y) \in M$ по некоторому правилу (закону) ставится в соответствие единственное значение переменной z.

Множество M упорядоченных пар чисел $(x; y)$ называется *областью определения* функции z, а множество всех возможных значений переменной z — *множеством значений функции*.

Тот факт, что переменная z является функцией двух переменных x и y, принято записывать так:

$$z = f(x, y).$$

В то время как для числовой функции одной переменной областью определения является множество точек числовой оси, область определения функции двух переменных представляет собой некоторое множество точек числовой плоскости.

Так же как и в случае функции одной переменной, для функций нескольких переменных существуют различные способы и формы их

задания. Так, при аналитическом способе задания функция нескольких независимых переменных может быть задана в явной, неявной или параметрической формах.

Примеры. 1. Формула

$$z = x^2 + y^3$$

определяет z как функцию двух независимых переменных x, y, заданную в явной форме. Область определения этой функции — множество всех упорядоченных пар действительных чисел $(x; y)$.

2. Равенство

$$e^z = x^2 - y^3$$

определяет z как функцию независимых переменных x, y, заданную в неявной форме. Область определения этой функции — множество всех упорядоченных пар действительных чисел $(x; y)$, удовлетворяющих неравенству $|x| > |y|$.

Подобно тому как функция $y = f(x)$ геометрически изображалась своим графиком в прямоугольной декартовой системе координат Oxy, функция двух переменных $z = f(x, y)$ в прямоугольной пространственной декартовой системе координат $Oxyz$ задает множество точек, которое является своего рода пространственным графиком функции $z = f(x, y)$. Пространственный график функции $z = f(x, y)$ обычно называют *поверхностью*, а равенство $z = f(x, y)$ — *уравнением поверхности*.

Простейшим примером такой поверхности является плоскость в пространстве, которая в прямоугольной декартовой системе координат определяется уравнением

$$Ax + By + Cz + D = 0.$$

Если $C \neq 0$, то уравнение может быть записано в виде

$$z = -\frac{A}{C}x - \frac{B}{C}y - \frac{D}{C}.$$

Аналогичным образом может быть введено понятие функции трех, четырех и т. д. переменных.

Все основные понятия теории функций одной переменной — такие, как предел функции, непрерывность функции, дифференцируемость и т. д. — обобщаются и на случай функций нескольких независимых переменных.

§ 6. Предел функции

6.1. Понятие предела функции. Пусть функция $f(x)$ определена во всех точках промежутка $(a; b)$, за исключением, быть может, некоторой точки $x_0 \in (a, b)$. Построим последовательность значений аргумента функции $f(x)$:

$$x_1, x_2, \ldots, x_n, \ldots, n \in \mathbf{N} \qquad (x_n \neq x_0), \tag{1}$$

такую, чтобы все члены последовательности принадлежали промежутку $(a; b)$ и последовательность сходилась к точке x_0:

$$\lim_{n \to \infty} x_n = x_0.$$

Тогда значения функции $f(x)$

$$f(x_1), f(x_2), \ldots, f(x_n), \ldots \tag{2}$$

также образуют некоторую числовую последовательность.

Говорят, что число A является *пределом функции* $f(x)$ при x, стремящемся к x_0, если для любой последовательности значений аргумента (1), сходящейся к числу x_0, последовательность значений функции (2) сходится к числу A, и пишут

$$\lim_{x \to x_0} f(x) = A.$$

Это определение предела функции называют определением предела по Гейне.

Существует и другое, эквивалентное сформулированному выше определение предела функции.

Говорят, что число A является *пределом функции* $f(x)$ при x, стремящемся к x_0, если для любого положительного числа ε найдется такое положительное число δ, зависящее от ε, что при всех $x \in (a; b)$, удовлетворяющих неравенству

$$0 < |x - x_0| < \delta,$$

выполняется неравенство

$$|f(x) - A| < \varepsilon.$$

Это определение предела функции называют определением предела по Коши.

Пример 1. Доказать, что предел функции $f(x) = \sqrt{x}$ при x, стремящемся к a ($a > 0$), равен $\sqrt{a}$.

Доказательство будем проводить, используя определение предела по Коши, т. е. покажем, что для любого числа $\varepsilon > 0$ найдется такое число $\delta > 0$ (для каждого ε, возможно, свое), что для всех x, принадлежащих области определения функции, из неравенства $0 < |x - a| < \delta$ следует неравенство $|\sqrt{x} - \sqrt{a}| < \varepsilon$.

Выражение $|\sqrt{x} - \sqrt{a}|$ можно записать в виде

$$|\sqrt{x} - \sqrt{a}| = \left|\frac{x-a}{\sqrt{x}+\sqrt{a}}\right| = \frac{|x-a|}{|\sqrt{x}+\sqrt{a}|} = \frac{|x-a|}{\sqrt{x}+\sqrt{a}}.$$

Отбрасывая в знаменателе последней дроби неотрицательное слагаемое $\sqrt{x}$ и тем самым увеличивая дробь, получим

$$|\sqrt{x} - \sqrt{a}| < \frac{|x-a|}{\sqrt{a}}. \tag{3}$$

Если взять $\delta = \varepsilon\sqrt{a}$, то из неравенства $0 < |x-a| < \delta$ будет следовать неравенство

$$\frac{|x-a|}{\sqrt{a}} < \frac{\delta}{\sqrt{a}} \Rightarrow \frac{|x-a|}{\sqrt{a}} < \varepsilon,$$

с учетом которого из (3) получаем $|\sqrt{x} - \sqrt{a}| < \varepsilon$.

Таким образом, для любого $\varepsilon > 0$ найдено такое $\delta > 0$ (а именно, $\delta = \varepsilon\sqrt{a}$), что для x, удовлетворяющих неравенству $0 < |x-a| < \delta$, будет выполнено неравенство $|\sqrt{x} - \sqrt{a}| < \varepsilon$.

Однако при таком выборе δ может оказаться, что среди значений x, удовлетворяющих неравенству $|x-a| < \delta$, есть значения x (а именно, отрицательные значения), которые не входят в область определения функции $f(x) = \sqrt{x}$.

Если выбрать в качестве δ минимальное из двух чисел $\varepsilon\sqrt{a}$ и a, то все значения x, удовлетворяющие неравенству

$$|x-a| < \delta,$$

принадлежат области определения функции $f(x)=\sqrt{x}$ и для них выполнено неравенство

$$|\sqrt{x}-\sqrt{a}|<\varepsilon,$$

т. е. число $\sqrt{a}$ действительно является пределом функции $f(x)=\sqrt{x}$ при x, стремящемся к a.

Пример 2. Доказать, что предел функции

$$f(x)=\frac{3x^2-3}{x+1}$$

при x, стремящемся к -1, равен -6. (Область определения данной функции: $(-\infty;-1)\cup(-1,+\infty)$.)

Докажем, что для любого числа $\varepsilon>0$ существует такое число $\delta=\delta(\varepsilon)>0$, что при всех $x\neq -1$, удовлетворяющих неравенству $|x+1|<\delta$, выполняется неравенство

$$\left|\frac{3x^2-3}{x+1}+6\right|<\varepsilon.$$

В самом деле, при $x\neq -1$ имеем

$$\left|\frac{3x^2-3}{x+1}+6\right|=\left|\frac{3(x+1)(x-1)}{x+1}+6\right|=\left|\frac{3(x-1)}{1}+6\right|=3|x+1|. \qquad (4)$$

Если взять δ равным, например, $\varepsilon/3$, то из неравенства $|x+1|<\delta=\varepsilon/3$ следует, что $3|x+1|<\varepsilon$, и в силу равенств (4) неравенство

$$\left|\frac{3x^2-3}{x+1}+6\right|<\varepsilon,$$

что и требовалось доказать.

Пример 3. Доказать, что функция

$$f(x)=\sin\frac{1}{x}$$

не имеет предела при x, стремящемся к нулю.

Доказательство будем проводить, используя определение предела по Гейне. Построим две последовательности значений аргумента:

$$\frac{2}{5\pi},\ \frac{2}{9\pi},\ \ldots,\ \frac{2}{\pi(1+4n)},\ \ldots;\qquad \frac{2}{7\pi},\ \frac{2}{11\pi},\ \ldots,\ \frac{2}{\pi(3+4n)},\ \ldots \qquad (5)$$

Обе последовательности сходятся к числу 0. Последовательности значений функции

$$f(x)=\sin\frac{1}{x},$$

соответствующие построенным последовательностям (5), следующие:

$$1,\ 1,\ 1,\ \ldots,\ 1,\ \ldots,$$
$$-1,\ -1,\ -1,\ \ldots,\ -1,\ \ldots$$

Пределами этих последовательностей будут числа 1 и -1 соответственно.

Таким образом, построены две последовательности значений аргумента функции

$$f(x)=\sin\frac{1}{x},$$

сходящиеся к одному и тому же числу, такие, что соответствующие последовательности значений данной функции имеют различные пределы; следовательно, данная функция не имеет предела при x, стремящемся к нулю. График функции $y=\sin\frac{1}{x}$ изображен на рис. 8.3.

Сформулированные выше определения предела функции по Гейне и по Коши могут быть обобщены и на случай, когда вместо числа x_0 берется $+\infty$ (или $-\infty$).

Так, например, пусть функция $f(x)$ определена на промежутке $(a; +\infty)$. Говорят, что число A является *пределом* функции $f(x)$ при x, стремящемся к $+\infty$, если для любого положительного числа ε найдется такое положительное число Δ, что для всех значений x, удовлетворяющих неравенству $x > \Delta$, выполняется неравенство

$$|f(x) - A| < \varepsilon;$$

в этом случае пушут

$$\lim_{x \to +\infty} f(x) = A.$$

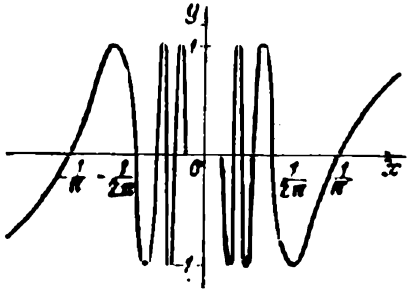


Рис. 8.5

Говорят, что функция $f(x)$ *стремится* к $+\infty$ при x, стремящемся к x_0, если для любого сколь угодно большого положительного числа E найдется такое число $\delta > 0$, что для всех значений x, удовлетворяющих неравенству

$$0 < |x - x_0| < \delta$$

и принадлежащих области определения функции, выполняется неравенство

$$f(x) > E;$$

в этом случае пишут

$$f(x) \to +\infty \quad \text{при} \quad x \to x_0.$$

Реже также используется символическая запись

$$\lim_{x \to x_0} f(x) = +\infty.$$

6.2. Односторонние пределы. Пусть функция $f(x)$ определена на промежутке $(a; x_0)$. Говорят, что число A является *левым пределом функции* $f(x)$ в точке x_0 и пишут

$$\lim_{x \to x_0 - 0} f(x) = A,$$

если для любого числа $\varepsilon > 0$ найдется положительное число δ, зависящее от ε, такое, что для всех $x \in (a; x_0)$, удовлетворяющих неравенству $x_0 - x < \delta$, выполняется неравенство $|f(x) - A| < \varepsilon$.

Аналогично определяется *правый предел функции* $f(x)$. Для обозначения правого предела функции $f(x)$ в точке x_0 используется запись

$$\lim_{x \to x_0 + 0} f(x).$$

Левый и правый пределы функции называются *односторонними пределами*.

Если функция $f(x)$ определена на промежутке $(a; b)$, за исключением, быть может, точки $x_0 \in (a; b)$, то для существования предела

$$\lim_{x \to x_0} f(x) = A$$

необходимо и достаточно, чтобы правый и левый пределы функции $f(x)$ в точке x_0 существовали и были равны:

$$\lim_{x \to x_0 + 0} f(x) = \lim_{x \to x_0 - 0} f(x) = A$$

В общем случае левый и правый пределы функции $f(x)$ в точке x_0 могут существовать, но не быть равными.

Так, например, пусть

$$f(x) = \operatorname{arctg} \frac{1}{x}.$$

Область определения данной функции — любое действительное число, отличное от нуля. Правый и левый пределы данной функции при x, стремящемся к нулю, соответственно равны

$$\lim_{x \to +0} \operatorname{arctg} \frac{1}{x} = \frac{\pi}{2},$$

$$\lim_{x \to -0} \operatorname{arctg} \frac{1}{x} = -\frac{\pi}{2},$$

т. е. правый и левый пределы существуют, но не равны. График этой функции изображен на рис. 8.6.

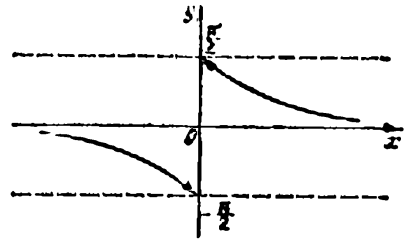


Рис. 8.6

6.3. Теоремы о пределах функций. 1. Если функции $f(x)$ и $g(x)$ имеют пределы при x, стремящемся к a, то функции $f(x) \pm g(x)$, $f(x) \cdot g(x)$, $\dfrac{f(x)}{g(x)}$ также имеют пределы при x, стремящемся к a, и

$$\lim_{x \to a} [f(x) \pm g(x)] = \lim_{x \to a} f(x) \pm \lim_{x \to a} g(x),$$

$$\lim_{x \to a} [f(x) \cdot g(x)] = \lim_{x \to a} f(x) \cdot \lim_{x \to a} g(x),$$

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x \to a} f(x)}{\lim\limits_{x \to a} g(x)}.$$

(В последнем случае предполагается, что функция $g(x)$ не обращается в нуль в достаточно малой окрестности точки a и $\lim\limits_{x \to a} g(x) \neq 0$.)

2. Если при x, стремящемся к a, функция $f(x)$ имеет предел, равный A, и этот предел больше числа c, то для достаточно близких к a значений x функция $f(x)$ удовлетворяет неравенству $f(x) > c$.

Использование теорем о пределах функций в ряде случаев позволяет свести вычисление предела функции к вычислению пределов более простых функций.

Пример 1. Вычислить предел

$$\lim_{x \to 4} \frac{\sqrt{1+2x} - 3}{\sqrt{x} - 2} \tag{6}$$

Непосредственное применение теоремы о пределе частного здесь невозможно, так как предел выражения, стоящего в знаменателе дроби, при $x \to 4$ равен нулю. Поэтому, прежде чем воспользоваться теоремой о пределе частного, необходимо так преобразовать данную дробь, чтобы ее знаменателем оказалось выражение, имеющее отличный от нуля предел. Для этого умножим числитель

и знаменатель выражения, стоящего под знаком предела, на выражение $\sqrt{1+2x}+3$:

$$\frac{\sqrt{1+2x}-3}{\sqrt{x}-2}=\frac{(\sqrt{1+2x}-3)(\sqrt{1+2x}+3)}{(\sqrt{x}-2)(\sqrt{1+2x}+3)}=\frac{2x-8}{(\sqrt{x}-2)(\sqrt{1+2x}+3)}=$$

$$=\frac{2(x-4)}{(\sqrt{x}-2)(\sqrt{1+2x}+3)}=\frac{2(\sqrt{x}+2)(\sqrt{x}-2)}{(\sqrt{x}-2)(\sqrt{1+2x}+3)}=\frac{2(\sqrt{x}+2)}{\sqrt{1+2x}+3}.$$

Предел функции, стоящей в знаменателе последней дроби, при $x\to 4$ существует и отличен от нуля, а предел функции, стоящей в числителе дроби, также существует. Поэтому к последнему выражению уже применима теорема о пределе частного:

$$\lim_{x\to 4}\frac{\sqrt{1+2x}-3}{\sqrt{x}-2}=\lim_{x\to 4}\frac{2(\sqrt{x}+2)}{\sqrt{1+2x}+3}=\frac{\lim\limits_{x\to 4}2(\sqrt{x}+2)}{\lim\limits_{x\to 4}(\sqrt{1+2x}+3)}.$$

Используя далее теоремы о пределах суммы и произведения функций, получим

$$\frac{\lim\limits_{x\to 4}(2\sqrt{x}+4)}{\lim\limits_{x\to 4}(\sqrt{1+2x}+3)}=\frac{2\lim\limits_{x\to 4}\sqrt{x}+\lim\limits_{x\to 4}4}{\lim\limits_{x\to 4}\sqrt{1+2x}+\lim\limits_{x\to 4}3}.$$

Принимая теперь во внимание, что $\lim\limits_{x\to a}\sqrt{x}=\sqrt{a}$ $(a>0)$, имеем

$$\frac{2\lim\limits_{x\to 4}\sqrt{x}+\lim\limits_{x\to 4}4}{\lim\limits_{x\to 4}\sqrt{1+2x}+\lim\limits_{x\to 4}3}=\frac{4+4}{3+3}=\frac{4}{3}.$$

Таким образом,

$$\lim_{x\to 4}\frac{\sqrt{1+2x}-3}{\sqrt{x}-2}=\frac{4}{3}.$$

Заметим, что здесь, так же как и при вычислении пределов последовательностей (см. п. 2.3), правомерность соединения знаком равенства выражений, получаемых в результате использования теорем о пределах, оказывается обоснованной лишь после того, как доказано, что исходная функция имеет предел.

Пример 2. Вычислить предел

$$\lim_{x\to+\infty}\frac{2x^2-x+1}{x^2-3}.$$

Так же как и в предыдущем примере, в данной задаче непосредственное использование теоремы о пределе частного невозможно, так как не существует как предела числителя, так и предела знаменателя дроби. Для того чтобы использовать теорему о пределе частного, преобразуем дробь, стоящую под знаком предела, поделив числитель и знаменатель дроби на старшую степень многочлена, стоящего в знаменателе дроби (а именно, на x^2):

$$\frac{2x^2-x+1}{x^2-3}=\frac{2-\frac{1}{x}+\frac{1}{x^2}}{1-\frac{3}{x^2}}.$$

Предел функции, стоящей в знаменателе дроби при $x\to+\infty$, существует и отличен от нуля, и предел функции, стоящей в числителе, также существует,

Поэтому по теореме о пределе частного имеем

$$\lim_{x\to+\infty}\frac{2x^2-x+1}{x^2-3}=\frac{\lim\limits_{x\to+\infty}\left(2-\frac{1}{x}+\frac{1}{x^2}\right)}{\lim\limits_{x\to+\infty}\left(1-\frac{3}{x^2}\right)}.$$

Используя далее теорему о пределе суммы и равенства

$$\lim_{x\to+\infty}\frac{1}{x}=0,\qquad \lim_{x\to+\infty}\frac{1}{x^2}=0,$$

справедливость которых нетрудно доказать исходя из определения предела функции, окончательно получаем

$$\lim_{x\to+\infty}\frac{2x^2-x+1}{x^2-3}=2.$$

В тех случаях, когда предел функции неизвестен и не может быть вычислен с помощью теорем о пределах, существование предела функции в ряде случаев может быть установлено с помощью необходимого и достаточного условия существования предела (критерия Коши).

Для того чтобы функция $f(x)$ при x, стремящемся к a, имела конечный предел, необходимо и достаточно, чтобы для любого $\varepsilon>0$ существовало положительное число δ (вообще говоря, зависящее от ε) такое, что для всех значений аргумента x' и x'', принадлежащих области определения и удовлетворяющих неравенствам

$$0<|x'-a|<\delta,\qquad 0<|x''-a|<\delta,$$

выполняется неравенство

$$|f(x')-f(x'')|<\varepsilon.$$

6.4. Пределы некоторых функций. В теории пределов важное место занимают перечисленные ниже пределы 1)—4), с помощью которых вычисляются многие пределы элементарных функций:

1) $\lim\limits_{x\to0}\dfrac{\sin x}{x}=1;$

2) $\lim\limits_{x\to0}(1+x)^{1/x}=e;$

3) $\lim\limits_{x\to0}\dfrac{a^x-1}{x}=\ln a,\qquad \lim\limits_{x\to0}\dfrac{e^x-1}{x}=1;$

4) $\lim\limits_{x\to0}\dfrac{\log_a(1+x)}{x}=\log_a e,\qquad \lim\limits_{x\to0}\dfrac{\ln(1+x)}{x}=1.$

Пример 1. Вычислить предел

$$\lim_{x\to0}\frac{1-\cos x}{x^2}.$$

Преобразуем дробь, стоящую под знаком предела:

$$\frac{1-\cos x}{x^2}=\frac{2\sin^2\frac{x}{2}}{x^2}=\frac{1}{2}\cdot\frac{\sin^2\frac{x}{2}}{\left(\frac{x}{2}\right)^2}.$$

Воспользовавшись пределом 1), получаем

$$\lim_{x\to0}\frac{1-\cos x}{x^2}=\frac{1}{2}\lim_{x\to0}\frac{\sin^2\frac{x}{2}}{\left(\frac{x}{2}\right)^2}=\frac{1}{2}.$$

С помощью предела 1) можно также получить ряд геометрических формул для вычислений длин, площадей и объемов.

Пример 2. Найти формулу для вычисления площади круга радиуса R.

По определению за площадь круга принимается предел последовательности площадей правильных многоугольников, вписанных в окружность при условии, что число сторон правильных многоугольников неограниченно возрастает.

Пусть $A_1, A_2, \ldots, A_n$ – вершины правильного n-треугольника, вписанного в окружность радиуса R с центром в точке O (на рис. 8.7 – шестиугольника

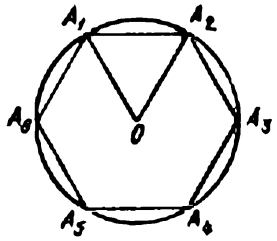


Рис. 8.7

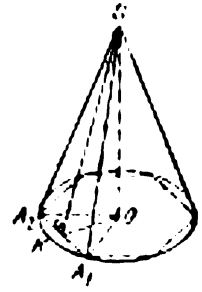


Рис. 8.8

$A_1A_2A_3A_4A_5A_6$). Рассмотрим треугольник A_1A_2O. Угол A_1OA_2 равен $2\pi/n$, а его площадь

$$S_{\triangle A_1OA_2}=\frac{1}{2}R^2\sin\frac{2\pi}{n}.$$

Площадь S_n правильного n-угольника, вписанного в окружность, равна

$$S_n=\frac{1}{2}R^2n\sin\frac{2\pi}{n}.$$

Рассмотрим последовательность площадей правильных n-угольников, вписанных в окружность, при условии, что число сторон неограниченно возрастает. Эта последовательность монотонно возрастающая и ограничена сверху (например, площадью любого правильного многоугольника, описанного вокруг окружности). Следовательно, по теореме Вейерштрасса (см. п. 2.4) эта последовательность имеет предел. Вычислим этот предел:

$$\lim_{n\to\infty}S_n=\lim_{n\to\infty}\left(\frac{1}{2}R^2n\sin\frac{2\pi}{n}\right)=\frac{1}{2}R^2\lim_{n\to\infty}\left(n\sin\frac{2\pi}{n}\right).$$

Записывая последний предел в виде

$$\frac{1}{2}R^2\lim_{n\to\infty}\frac{\sin\frac{2\pi}{n}}{\frac{1}{n}}=\frac{1}{2}R^2\cdot2\pi\lim_{n\to\infty}\left(\frac{\sin\frac{2\pi}{n}}{\frac{2\pi}{n}}\right)$$

и обозначая $\frac{2\pi}{n}=x$, нетрудно заметить, что вычисление исходного предела свелось к вычислению предела

$$\lim_{x\to0}\frac{\sin x}{x},$$

который равен единице.

Таким образом, окончательно имеем

$$\lim_{n\to\infty}S_n=\pi R^2\lim_{x\to0}\frac{\sin x}{x}=\pi R^2.$$

Аналогичным образом могут быть получены формула для вычисления длины окружности, формулы для вычисления площадей боковой поверхности цилиндра и конуса, а также формулы для вычисления объемов этих фигур.

Пример 3. Вывести формулу для вычисления площади боковой поверхности конуса.

По определению за площадь боковой поверхности конуса принимается предел последовательности площадей боковых поверхностей правильных пирамид, вписанных в конус, при условии, что число сторон правильных многоугольников, лежащих в основании пирамид, неограниченно возрастает.

Обозначим высоту конуса через H, а радиус окружности, лежащей в основании конуса, — через R. Пусть $A_1, A_2, \ldots, A_n$ — вершины правильного n-угольника, вписанного в окружность основания конуса (на рис. 88 — шестиугольника). Рассмотрим треугольник A_1OA_2 (O — центр окружности). Угол A_1OA_2 равен $\frac{2\pi}{n}$; $|A_1A_2| = 2R\sin\frac{\pi}{n}$. Высота треугольника $|OK| = R\cos\frac{\pi}{n}$. Апофему пирамиды $SA_1A_2 \ldots A_n$ вычислим по теореме Пифагора из треугольника SOK:

$$|SK| = \sqrt{|SO|^2 + |OK|^2} = \sqrt{H^2 + R^2\cos^2\frac{\pi}{n}},$$

Периметр правильного многоугольника $A_1A_2 \ldots A_n$ равен

$$P_n = 2Rn\sin\frac{\pi}{n},$$

а площадь боковой поверхности правильной n-угольной пирамиды равна

$$S_n = \frac{1}{2}|SK| \cdot P_n = \sqrt{H^2 + R^2\cos^2\frac{\pi}{n}} \cdot Rn\sin\frac{\pi}{n}.$$

Вычислим предел, который принимается за площадь боковой поверхности конуса:

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty}\left(\sqrt{H^2 + R^2\cos^2\frac{\pi}{n}} \cdot Rn\sin\frac{\pi}{n}\right) =$$

$$= \lim_{n\to\infty}\sqrt{H^2 + R^2\cos^2\frac{\pi}{n}} \cdot \lim_{n\to\infty} Rn\sin\frac{\pi}{n} =$$

$$= \sqrt{H^2 + R^2}\,\pi R \lim_{n\to\infty}\frac{\sin\frac{\pi}{n}}{\frac{\pi}{n}} = \pi R\sqrt{H^2 + R^2}.$$

Поскольку величина $\sqrt{H^2 + R^2} = L$ — образующая конуса, то формула для вычисления площади боковой поверхности конуса принимает вид

$$S_{\text{бок}} = \lim_{n\to\infty} S_n = \pi RL.$$

§ 7. Бесконечно малые и бесконечно большие величины

7.1. Понятие бесконечно малой и бесконечно большой величины. Числовая функция $f(x)$ называется *бесконечно малой функцией* (или *бесконечно малой величиной*) при x, стремящемся к a, если

$$\lim_{x\to a} f(x) = 0.$$

Иными словами, функция $f(x)$ называется *бесконечно малой* при x, стремящемся к a, если для любого числа $\varepsilon > 0$ найдется такое положительное число δ, зависящее от ε, что при всех $x \neq a$, принадлежащих области определения функции и удовлетворяющих неравенству

$$|x - a| < \delta,$$

выполняется неравенство

$$|f(x)| < \varepsilon.$$

Подчеркнем, что бесконечно малую величину следует понимать как переменную величину, которая лишь в процессе своего изменения (при стремлении x к a) делается меньше произвольного числа ε. Поэтому, например, утверждение типа «одна миллионная есть бесконечно малая величина» неверно: о числе не имеет смысла говорить, что оно бесконечно малое.

Примеры бесконечно малых величин:

$f(x)=x^2$ при x, стремящемся к нулю;
$f(x)=\cos x$ при x, стремящемся к $\pi/2$;
$f(x)=\ln x$ при x, стремящемся к единице.

Функция $f(x)$ называется *бесконечно большой величиной* при x, стремящемся к a, если для любого сколь угодно большого числа $E > 0$ найдется такое положительное число δ, зависящее от E, что при всех $x \neq a$, принадлежащих области определения функции и удовлетворяющих неравенству

$$0 < |x-a| < \delta(E),$$

выполняется неравенство

$$|f(x)| > E.$$

Как и в случае бесконечно малых, ни одно отдельно взятое значение бесконечно большой величины не может быть названо как «бесконечно большое»; бесконечно большая величина — это функция, которая лишь в процессе своего изменения может стать больше произвольно взятого числа E.

Примеры бесконечно больших величин:

$f(x)=\frac{1}{x}$ при x, стремящемся к нулю;
$f(x)=\operatorname{tg} x$ при x, стремящемся к $\pi/2$;
$f(x)=\ln|x|$ при x, стремящемся к нулю;
$f(x)=\frac{1}{(x-1)^2}$ при x, стремящемся к единице.

Если функция $f(x)$ при x, стремящемся к a, бесконечно большая и для всех достаточно близких к a значений x сохраняет определенный знак (плюс или минус), то в соответствии со знаком говорят, что функция $f(x)$ стремится к $+\infty$ $(-\infty)$, и пишут

$$f(x) \longrightarrow +\infty \quad \text{при} \quad x \longrightarrow a \quad \text{или} \quad \lim_{x \to a} f(x) = +\infty$$

$$\left(f(x) \longrightarrow -\infty \quad \text{при} \quad x \longrightarrow a \quad \text{или} \quad \lim_{x \to a} f(x) = -\infty\right).$$

Из приведенных выше примеров бесконечно больших:

$$\lim_{x \to 0} \ln|x| = -\infty, \quad \lim_{x \to 1} \frac{1}{(x-1)^2} = +\infty.$$

Бесконечно малые и бесконечно большие величины связаны между собой следующей зависимостью.

Если $f(x)$ — бесконечно большая функция (величина) при x, стремящемся к a, то ее обратная величина

$$\alpha(x) = \frac{1}{f(x)}$$

— бесконечно малая функция (величина) при x, стремящемся к a.

7.2. Сравнение бесконечно малых и бесконечно больших величин. Приведенные в п. 7.1 примеры различных бесконечно малых наводят на мысль о том, не будут ли одни бесконечно малые стремиться к нулю «быстрее», чем другие? Прежде всего определим, что следует понимать под словом «быстрее», т. е. как сравнивать две бесконечно малые величины.

Две бесконечно малые величины $\alpha(x)$ и $\beta(x)$ при x, стремящемся к a, сравниваются следующим образом. Рассмотрим отношение этих величин

$$\frac{\beta(x)}{\alpha(x)}$$

(при этом полагается, что величина, стоящая в знаменателе дроби, отлична от нуля по крайней мере для значений x, достаточно близких к числу a) и вычислим предел этого отношения при x, стремящемся к a.

1. Если отношение $\frac{\beta(x)}{\alpha(x)}$ имеет конечный, отличный от нуля предел при x, стремящемся к a, то бесконечно малые $\alpha(x)$ и $\beta(x)$ считаются *бесконечно малыми одного порядка.*

2. Если при x, стремящемся к a, ни одно из двух отношений $\frac{\alpha(x)}{\beta(x)}$, $\frac{\beta(x)}{\alpha(x)}$ предела не имеет, то бесконечно малые величины $\alpha(x)$ и $\beta(x)$ считаются *не сравнимыми* между собой.

3. Если отношение $\frac{\beta(x)}{\alpha(x)}$ само оказывается бесконечно малым $\left(\text{а отношение } \frac{\alpha(x)}{\beta(x)} \text{ — бесконечно большим}\right)$ при x, стремящемся к a, то бесконечно малая $\beta(x)$ считается величиной *высшего порядка малости* по сравнению с бесконечно малой $\alpha(x)$, а бесконечно малая $\alpha(x)$ считается величиной *низшего порядка малости* по сравнению с величиной $\beta(x)$.

Например, при x, стремящемся к нулю, бесконечно малыми величинами одного порядка будут следующие пары функций:

$$\alpha(x)=x, \quad \beta(x)=\sqrt{1+x}-1;$$
$$\alpha(x)=x, \quad \beta(x)=\sin x.$$

При x, стремящемся к нулю, бесконечно малая $\beta(x)=1-\cos x$ является бесконечно малой высшего порядка по сравнению с бесконечно малой $\alpha(x)=x$, так как

$$\lim_{x\to 0}\frac{\beta(x)}{\alpha(x)}=\lim_{x\to 0}\frac{1-\cos x}{x}=\lim_{x\to 0}\frac{2\sin^2\frac{x}{2}}{x}=$$
$$=\lim_{x\to 0}\frac{\sin^2\frac{x}{2}}{\frac{x}{2}}=\lim_{x\to 0}\frac{\sin\frac{x}{2}}{\frac{x}{2}}\cdot\lim_{x\to 0}\sin\frac{x}{2}=1\cdot 0=0.$$

Бесконечно малые величины $\alpha(x)=x$ и $\beta(x)=x\sin\frac{1}{x}$ при x, стремящемся к нулю, не сравнимы между собой.

Бесконечно малая величина высшего порядка по сравнению с бесконечно малой величиной $\alpha(x)$ обычно обозначается символом $o(\alpha)$.

Бесконечно малую $\beta(x)$ называют *бесконечно малой k-го порядка* по сравнению с величиной $\alpha(x)$, если величины $\beta(x)$ и $\alpha^k(x)$ являются величинами одного порядка.

Например, при x, стремящемся к нулю, бесконечно малая величина $\beta(x)=1-\cos x$ относительно бесконечно малой $\alpha(x)=x$ — бесконечно малая второго порядка, так как

$$\lim_{x\to 0}\frac{1-\cos x}{x^2}=\frac{1}{2}.$$

Если из нескольких бесконечно малых $\alpha(x)$, $\beta(x)$, $\gamma(x)$, ... выбрана одна бесконечно малая $\alpha(x)$ (*основная* бесконечно малая), то бесконечно малые вида $c\cdot\alpha^k$ (c — постоянная) называют *простейшими* бесконечно малыми.

Если бесконечно малая $\beta(x)$ — бесконечно малая k-го порядка относительно $\alpha(x)$, т. е.

$$\lim\frac{\beta(x)}{\alpha^k(x)}=c,\quad c\neq 0,\quad \text{или}\quad \lim\frac{\beta(x)}{c\alpha^k(x)}=1,$$

то бесконечно малые $c\cdot\alpha^k(x)$ и $\beta(x)$ называют *эквивалентными*.

Простейшая бесконечно малая $c\cdot\alpha^k$, эквивалентная бесконечно малой $\beta(x)$, называется ее *главной частью*.

Например, при $x\to 0$ простейшая бесконечно малая $\frac{1}{2}x^2$ является главной частью бесконечно малой $1-\cos x$.

Бесконечно большие величины сравниваются между собой так же, как и бесконечно малые величины.

Пусть $f(x)$ и $g(x)$ при x, стремящемся к a, — две бесконечно большие величины.

1. $f(x)$ и $g(x)$ считаются величинами *одного порядка* при x, стремящемся к a, если их отношение $f(x)/g(x)$ имеет конечный и отличный от нуля предел.

2. Если отношение $f(x)/g(x)$ при x, стремящемся к a, само оказывается бесконечно большим (а отношение $g(x)/f(x)$ бесконечно малым), то величина $f(x)$ считается *бесконечно большой высшего порядка* по сравнению с $g(x)$, а $g(x)$ — *бесконечно большой низшего порядка* по сравнению с $f(x)$.

3. Если отношение $f(x)/g(x)$ ни к какому конечному пределу не стремится и в то же время не является бесконечно большим, то $f(x)$ и $g(x)$ считаются *несравнимыми*.

§ 8. Непрерывность (и разрывы) функций

8.1. Понятие непрерывности функции. С понятием предела функции тесно связано другое важное понятие математического анализа — понятие непрерывности функции.

В определении предела функции $f(x)$ при x, стремящемся к x_0, отмечалось, что число x_0 может и не принадлежать области определения функции $f(x)$ (см. п. 7.1), а если число x_0 и принадлежит области определения, значение функции в этой точке при изучении предела нас не интересовало.

Рассмотрим функцию $f(x)$, определенную на некотором промежутке $(a;\ b)$. Пусть x_0 — некоторая точка промежутка $(a;\ b)$, в которой функция имеет значение $f(x_0)$.

Если

$$\lim_{x \to x_0} f(x) = f(x_0), \tag{1}$$

то говорят, что функция $f(x)$ *непрерывна* в точке x_0. Если же $\lim\limits_{x \to x_0} f(x)$ не существует либо равенство (1) не выполняется, то говорят, что в точке x_0 функция терпит *разрыв*.

Таким образом, выяснение вопроса о непрерывности функции $f(x)$ в какой-либо точке x_0, принадлежащей области определения функции, сводится к вычислению предела функции в точке x_0 и проверке справедливости равенства (1).

Если функция $f(x)$ определена на промежутке (a, b) и непрерывна в каждой точке промежутка, то говорят, что функция *непрерывна на промежутке*.

8.2. Основные теоремы о непрерывных функциях. Непрерывность элементарных функций.

1. Если функции $f(x)$ и $g(x)$ определены на промежутке (a, b) и непрерывны в точке $x_0 \in (a, b)$, то в этой точке функции

$$f(x) \pm g(x), \quad f(x) \cdot g(x), \quad \frac{f(x)}{g(x)}$$

также непрерывны (последняя функция непрерывна при условии, что $g(x_0) \neq 0$).

2. Если функция $y = f(x)$ непрерывна в точке x_0, а функция $g(y)$ непрерывна в точке $y_0 = f(x_0)$, то и сложная функция $g(f(x))$ непрерывна в точке x_0.

3. Первая теорема Больцано—Коши. Если функция $f(x)$ задана на промежутке $[a; b]$ и непрерывна в промежутке $[a; b]$ и если на концах промежутка значения функции $f(a)$ и $f(b)$ имеют разные знаки, то между a и b существует (по крайней мере одно) такое значение x_0, при котором $f(x)$ обращается в нуль:

$$f(x_0) = 0, \quad x_0 \in (a; b).$$

Теорема Больцано — Коши, в частности, позволяет доказать, что любое алгебраическое уравнение нечетной степени с действительными коэффициентами

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n = 0$$

имеет по крайней мере один действительный корень.

Действительно, при достаточно больших по абсолютной величине значениях x многочлен, стоящий в левой части уравнения, имеет знак старшего члена, т. е. при $x > 0$ — знак a_0, а при $x < 0$ — противоположный знак. Так как многочлен — непрерывная функция на всем множестве действительных чисел, то по теореме Больцано — Коши обязательно найдется такая точка x_0, в которой этот многочлен обратится в нуль, и, следовательно, алгебраическое уравнение обратится в тождество.

4. Первая теорема Вейерштрасса. Если функция $f(x)$ определена и непрерывна на промежутке $[a, b]$, то она ограничена на этом промежутке, т. е. существуют такие два числа m и M, что для всех $x \in [a; b]$ имеет место неравенство $m \leqslant f(x) \leqslant M$.

Используя определение непрерывности функции в точке, нетрудно доказать, что функции $f(x) = c$ и $f(x) = x$ непрерывны в любой точке числовой оси. Из непрерывности этих функций по теореме 1 следует непрерывность целой рациональной функции

$$f(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$$

Таблица 7

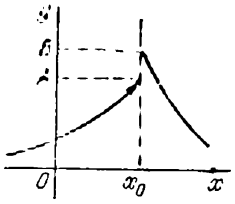


Рис. 8.9

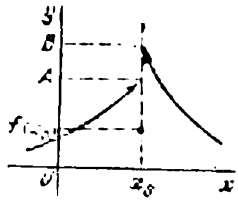


Рис. 8.10

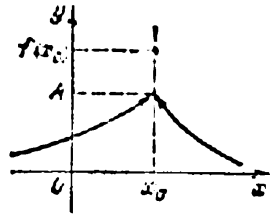


Рис. 8.11

Точка x_0 принадлежит области определения функции $f(x)$ — промежутку $(a; b)$:

$$x_0 \in (a; b)$$

Левый и правый пределы функции $f(x)$ в точке x_0 существуют и конечны (A и B — числа).

1. $\lim\limits_{x \to x_0 - 0} f(x) = A$, $\lim\limits_{x \to x_0 + 0} f(x) = B$, $A \neq B$, $f(x_0) = A$ или $f(x_0) = B$ (на рис. 8.9 $f(x_0) \neq B$). В этом случае говорят, что в точке x_0 функция $f(x)$ терпит скачок, равный $B - A$.

2. $\lim\limits_{x \to x_0 - 0} f(x) \neq A$, $\lim\limits_{x \to x_0 + 0} f(x) \neq B$, $A \neq B$, $f(x_0) \neq A$, $f(x_0) \neq B$ (рис. 8.10).

3. $\lim\limits_{x \to x_0 - 0} f(x) = \lim\limits_{x \to x_0 + 0} f(x) = A$, $f(x_0) \neq A$ (рис. 8.11). Функция терпит скачок, равный $f(x_0) - A$.

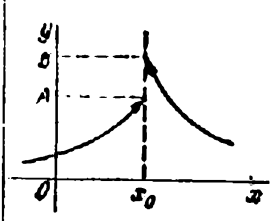


Рис. 8.12

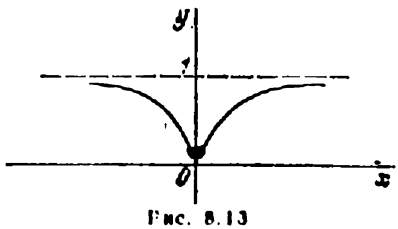


Рис. 8.13

В точке x_0 функция $f(x)$ — не определена и область определения функции

$$D(f) = (a; x_0) \cup (x_0; b)$$

Левый и правый пределы функции $f(x)$ в точке x_0 существуют и конечны.

1. $\lim\limits_{x \to x_0 - 0} f(x) = A$, $\lim\limits_{x \to x_0 + 0} f(x) = B$, $A \neq B$. Функция терпит скачок, равный $B - A$ (рис. 8.12).

2. $\lim\limits_{x \to x_0 - 0} f(x) = \lim\limits_{x \to x_0 + 0} f(x) = A$. В этом случае говорят, что точка x_0 — точка устранимого разрыва, так как, доопределяя функцию $f(x)$ в точке x_0 равенством $f(x_0) = A$, получаем непрерывную функцию. Примером такой функции может служить функция $f(x) = 2^{-\frac{1}{x^2}}$, не определенная в точке $x = 0$ (рис. 8.13). Для этой функции в точке $x = 0$ имеем

$$\lim_{x \to -0} 2^{-\frac{1}{x^2}} = \lim_{x \to +0} 2^{-\frac{1}{x^2}} = 0.$$

Доопределяя данную функцию равенством $f(0) = 0$, получаем непрерывную функцию в точке $x = 0$.

и непрерывность рациональной функции

$$f(x)=\frac{a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n}{b_0x^m+b_1x^{m-1}+\ldots+b_{m-1}x+b_m}$$

при всех значениях x, не обращающих знаменатель в нуль.

Используя определение (1) п. 8.1, можно доказать, что функции $\sin x$ и $\cos x$ непрерывны в любой точке числовой оси. Из непрерывности этих функций в силу теоремы 1 следует, что функции $\operatorname{tg} x$, $\operatorname{ctg} x$, $\sec x$ и $\operatorname{cosec} x$ также непрерывны при всех значениях аргумента, принадлежащих области определения функций.

Непрерывность обратных тригонометрических функций следует из свойств обратных функций (см. п. 5.9).

Степенная функция x^α, показательная функция a^x и логарифмическая функция $\log_a x$ непрерывны в своей области определения.

8.3. Односторонняя непрерывность. Классификация разрывов функций. Пусть функция $f(x)$ определена на промежутке $(a; x_0]$. Говорят, что функция $f(x)$ непрерывна в точке x_0 *слева*, если

$$\lim_{x\to x_0-0} f(x)=f(x_0). \tag{2}$$

Аналогично функция $f(x)$, определенная на промежутке $[x_0; b)$, непрерывна в точке x_0 *справа*, если

$$\lim_{x\to x_0+0} f(x)=f(x_0). \tag{3}$$

Если равенство (2) не выполняется, то говорят, что функция $f(x)$ *разрывна в точке x_0 слева*; если не выполняется равенство (3), то функция *разрывна в точке x_0 справа*.

Если функция $f(x)$ определена на промежутке $(a; b)$ и точка $x_0\in(a; b)$, то для непрерывности функции $f(x)$ в точке x_0 необходимо и достаточно, чтобы функция $f(x)$ была непрерывна слева и справа в точке x_0.

Различают следующие виды разрывов функции в точке: разрывы первого рода и разрывы второго рода.

Перечень разрывов первого рода приведен в табл. 7.

Если хотя бы один из односторонних пределов в точке x_0 не существует или равен $+\infty(-\infty)$, то говорят, что функция $f(x)$ имеет в точке x_0 разрыв второго рода (точка x_0 может как принадлежать, так и не принадлежать области определения функции). Примером функции, имеющей разрыв второго рода, может служить функция

$$f(x)=2^{1/x},$$

не определенная в нуле, для которой

$$\lim_{x\to+0} 2^{1/x}=+\infty, \qquad \lim_{x\to-0} 2^{1/x}=0.$$

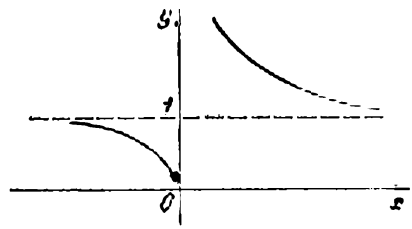


Рис. 8.14

График этой функции изображен на рис. 8.14.

8.4. Предел и непрерывность функций двух независимых переменных. Пусть $(x_0; y_0)$ — некоторая фиксированная точка координатной плоскости. Множество всех точек плоскости, координаты которых удовлетворяют неравенству

$$(x-x_0)^2+(y-y_0)^2<\varepsilon^2,$$

называют ε-*окрестностью* точки $M_0(x_0; y_0)$. Расстояние от любой точки ε-окрестности до точки $M_0(x_0; y_0)$ меньше ε.

Пусть D — некоторое множество точек координатной плоскости Oxy. Точку $M_0(x_0; y_0)$ называют *внутренней точкой* этого множества, если она принадлежит множеству D вместе с некоторой достаточно малой ее окрестностью.

Множество, состоящее только из внутренних точек, называют *открытой областью*.

Точка $M_0(x_0, y_0)$ называется точкой *сгущения* множества D, если в каждой ее ε-окрестности содержится хотя бы одна точка множества D, отличная от точки M_0.

Точки сгущения открытой области, не принадлежащие ей, называются *граничными* точками этой области. Множество всех граничных точек области D образуют *границу* области.

Открытая область вместе со своей границей называется *замкнутой* областью.

Так, например, границей открытой области (открытого круга), задаваемой неравенством $(x-x_0)^2+(y-y_0)^2 < R^2$, является окружность $(x-x_0)^2+(y-y_0)^2=R^2$.

Пусть $z=f(x, y)$ — некоторая функция двух независимых переменных, определенная на открытой области D, и $M(a; b)$ — точка этой области.

Говорят, что функция $f(x, y)$ имеет *пределом* число A при стремлении переменных x, y к числам a, b соответственно, если для любого сколь угодно малого положительного числа ε найдется такое число δ, что неравенство

$$|f(x, y) - A| < \varepsilon$$

выполняется для всех значений x, y, принадлежащих D и удовлетворяющих неравенствам

$$0 < |x-a| < \delta, \quad 0 < |y-b| < \delta.$$

На «геометрическом» языке данное определение можно сформулировать так.

Число A называют *пределом функции* $f(x, y)$ при стремлении точки $M(x; y)$ к точке $M_0(a; b)$ и пишут

$$\lim_{M \to M_0} f(x, y) = A \quad \text{или} \quad \lim_{\substack{x \to a \\ y \to b}} f(x, y) = A,$$

если для любого сколь угодно малого числа $\varepsilon > 0$ существует такое число $\delta > 0$, что неравенство

$$|f(x, y) - A| < \varepsilon$$

выполняется для всех точек $M \in D$, удовлетворяющих неравенству

$$0 < |MM_0| < \delta,$$

где $|MM_0| = \sqrt{(x-a)^2+(y-b)^2}$ — расстояние между точками $M(x; y)$ и $M_0(a; b)$.

Говорят, что функция $f(x, y)$ *непрерывна* в точке $M_0(a; b)$, если выполняется неравенство

$$\lim_{M \to M_0} f(x, y) = f(a, b).$$

В противном случае говорят, что функция $f(x, y)$ в точке M_0 имеет разрыв.

Из непрерывности функции двух переменных $f(x, y)$ отдельно по каждой переменной не следует, что функция будет непрерывна как функция двух переменных. Так, например, функция двух переменных $f(x, y)$, определенная на всей координатной плоскости и задаваемая условиями

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2+y^2}, & \text{если } x^2+y^2 \neq 0, \\ 0, & \text{если } x^2+y^2 = 0, \end{cases}$$

непрерывна по каждой переменной в отдельности: при любом фиксированном значении y функция $u(x)=f(x, y)$ непрерывна при любом значении x, и при любом фиксированном значении x функция $v(y)=f(x, y)$ непрерывна при любом значении y.

Однако функция $f(x, y)$ разрывна в начале координат, так как в любой сколь угодно малой ε-окрестности точки (0; 0) существуют точки с координатами $(a; a)$, в которых значение функции равно 1/2.

Глава 9

НАЧАЛА ДИФФЕРЕНЦИАЛЬНОГО И ИНТЕГРАЛЬНОГО ИСЧИСЛЕНИЙ

Математический анализ как раздел математики возник в результате объединения двух различных и первоначально не связанных направлений математических исследований — дифференциального и интегрального исчислений.

Первоначально интуитивное представление о математическом объекте, который мы сейчас называем определенным интегралом, встречалось в работах ученых Древней Греции. Так, Архимед для вычисления объемов и площадей поверхностей тел пользовался разбиением фигур на элементы с последующим суммированием этих элементов, предвосхищая тем самым понятия интегральных сумм. Аналогичными задачами, развивая метод Архимеда, занимались И. Кеплер, Б. Паскаль, П. Ферма и другие ученые. Ферма также занимался задачами, которые мы сейчас относим к дифференциальному исчислению,— проведением касательных к кривым, нахождением наибольшего и наименьшего значений функций и т. д., причем для решения этих задач он, по существу, пользовался понятием приращения функции. Связь между этими различными классами задач была осознана учеными после исследований И. Ньютона и Г. Лейбница. Лейбницем и были введены используемые в настоящее время обозначения интеграла и дифференциала. Строгое обоснование большинства понятий математического анализа было дано Коши в середине XIX столетия на основе теории пределов. Дальнейшее развитие математического анализа привело к выделению таких самостоятельных разделов математики, как теория обыкновенных дифференциальных уравнений, теория дифференциальных уравнений в частных производных, теория интегральных уравнений, теория функций комплексной переменной, теория функций действительной переменной, функционального анализа и т. д.

§ 1. Производная

1.1. Понятие производной. Пусть $f(x)$ — некоторая функция, определенная на промежутке $(a; b)$, и x_0 — некоторая фиксированная точка этого промежутка. Возьмем произвольное значение x из промежутка $(a; b)$ и составим разность

$$x - x_0.$$

Разность $x - x_0$ называют *приращением независимой переменной* (или *приращением аргумента*) функции $f(x)$ в точке x_0 и обозначают Δx:

$$x - x_0 = \Delta x. \tag{1}$$

Приращением функции $f(x)$ в точке x_0 называют разность между значением функции в точке $x = x_0 + \Delta x$ и значением функции в точке

x_0 и обозначают $\Delta f(x_0)$:

$$f(x_0+\Delta x)-f(x_0)=\Delta f(x_0), \tag{2}$$

Так как точка x_0 считается фиксированной, приращение функции $\Delta f(x_0)$ является функцией приращения аргумента Δx. Составим отношение

$$\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}=\frac{\Delta f(x_0)}{\Delta x},$$

которое также будет функцией приращения аргумента Δx, и рассмотрим предел этого выражения при Δx, стремящемся к нулю:

$$\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}.$$

Если этот предел существует, то говорят, что функция $f(x)$ *имеет производную* в точке x_0, и пишут

$$\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}=f'(x_0). \tag{3}$$

Число $f'(x_0)$ называется *производной* функции в точке x_0. Производную функции $f(x)$ в точке x_0 также обозначают

$$f'(x)\big|_{x=x_0} \quad \text{или} \quad \frac{df(x_0)}{dx}.$$

Если существует предел (3), также говорят, что функция $f(x)$ *дифференцируема* в точке x_0 *). Если функция $f(x)$ дифференцируема в каждой точке промежутка $(a;\ b)$, то говорят, что она *дифференцируема в промежутке* $(a;\ b)$. Производная функции $f(x)$, дифференцируемой в промежутке $(a;\ b)$, сама является функцией x.

Примеры. 1. Вычислить производную функции $f(x)=x^2$ в точке x_0 числовой оси.

Вычисляя приращение функции

$$f(x_0+\Delta x)-f(x_0)=(x_0+\Delta x)^2-x_0^2=2x_0\cdot\Delta x+(\Delta x)^2,$$

по определению производной имеем

$$\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{2x_0\,\Delta x+(\Delta x)^2}{\Delta x}=\lim_{\Delta x\to 0}(2x_0+\Delta x)=2x_0,$$

т. е. $f'(x_0)=(x^2)'\big|_{x=x_0}=2x_0$ для любой точки x_0.

2. Доказать, что производная функции $f(x)=|x|$ в точке $x=0$ не существует.

Составим отношение приращения функции к приращению аргумента:

$$\frac{\Delta f(0)}{\Delta x}=\frac{f(0+\Delta x)-f(0)}{\Delta x}=\frac{|\Delta x|}{\Delta x}.$$

По определению модуля

$$|\Delta x|=\begin{cases}\Delta x, & \text{если } \Delta x\geqslant 0,\\ -(\Delta x), & \text{если } \Delta x<0.\end{cases}$$

Вычислим правый и левый пределы отношения $\dfrac{\Delta f(0)}{\Delta x}$, учитывая, что при

*) О дифференцируемости функции см также п. 2.1.

вычислении правого предела $\Delta x > 0$, а при вычислении левого $\Delta x < 0$:

$$\lim_{\Delta x \to +0} \frac{|\Delta x|}{\Delta x} = \lim_{\Delta x \to +0} \frac{\Delta x}{\Delta x} = \lim_{\Delta x \to +0} 1 = 1,$$

$$\lim_{\Delta x \to -0} \frac{|\Delta x|}{\Delta x} = \lim_{\Delta x \to -0} \frac{-(\Delta x)}{\Delta x} = \lim_{\Delta x \to -0} (-1) = -1.$$

Так как правый предел отношения $\frac{\Delta |x|}{\Delta x}$ не равен левому пределу, то обычного (двустороннего) предела данного отношения не существует и, следовательно, производная функции $f(x) = |x|$ при $x = 0$ не существует.

1.2. Геометрическая интерпретация производной. Пусть дана некоторая линия L и на ней точка M (рис. 9.1). Возьмем на линии L некоторую точку N, не совпадающую с точкой M. Прямую MN называют *секущей* для линии L. Пусть теперь точка N приближается к точке M, оставаясь на линии L. Тогда каждому положению

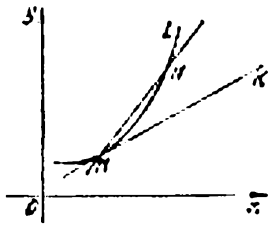


Рис. 9.1

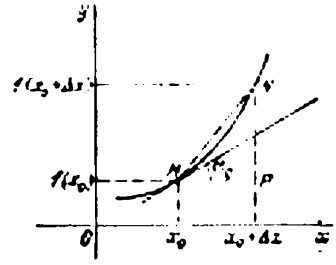


Рис. 9.2

точки N будет соответствовать своя секущая и все эти секущие будут проходить через точку M.

Касательной к линии L в точке M называется предельное положение MK секущей MN при стремлении точки N к точке M.

Пусть $y = f(x)$ — некоторая функция, дифференцируемая в точке x_0. В декартовой прямоугольной системе координат точка M, лежащая на графике функции $y = f(x)$ и имеющая абсциссу x_0, имеет координаты $(x_0; f(x_0))$. Пусть точка N принадлежит графику функции (рис. 9.2) и имеет координаты $(x_0 + \Delta x; f(x_0 + \Delta x))$. Проведем через точку M прямую, параллельную оси Ox, и обозначим точку пересечения этой прямой и прямой $x_0 + \Delta x$ через P. Рассмотрим прямоугольный треугольник MNP. Отношение

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

равно тангенсу угла наклона секущей MN к положительному направлению оси Ox:

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \operatorname{tg} \alpha. \tag{4}$$

Если устремить приращение Δx к нулю, то геометрически это будет означать, что точка N с координатами $(x_0 + \Delta x, f(x_0 + \Delta x))$ будет двигаться по графику функции $y = f(x)$, приближаясь к точке M, а угол α будет стремиться к углу α_0 — углу наклона касательной к положительному направлению оси Ox.

Так как предел левой части равенства (4) при $\Delta x \to 0$ равен $f'(x_0)$:

$$\lim_{\Delta x \to 0} \frac{f(x_0+\Delta x)-f(x_0)}{\Delta x} = f'(x_0),$$

а предел правой части равен $\operatorname{tg}\alpha_0$: $\lim\limits_{\Delta x \to 0} \operatorname{tg}\alpha = \operatorname{tg}\alpha_0$, то

$$f'(x_0) = \operatorname{tg}\alpha_0, \tag{5}$$

т. е. значение функции $f'(x)$ в точке x_0 равно тангенсу угла наклона касательной.

Используя формулу (5), уравнение касательной к графику функции $y = f(x)$, проходящей через точку с координатами $(x_0; y_0 = f(x_0))$, можно записать в виде

$$y - y_0 = f'(x_0)\cdot(x - x_0).$$

1.3. Кинематический смысл производной. Пусть материальная точка движется по прямой под действием некоторых сил не меняя направления своего движения, и пусть $S(t)$ — расстояние, пройденное точкой от некоторого момента времени, который принят за нулевой, до момента t. Выберем какой-либо момент времени t_0 и рассмотрим промежуток времени Δt от момента t_0 до момента $t = t_0 + \Delta t$. За этот промежуток времени точка пройдет некоторый путь, который обозначим $\Delta S(t_0)$. Этот путь есть функция Δt. По известному из физики определению отношение

$$\frac{\Delta S(t_0)}{\Delta t}$$

есть средняя скорость движения точки за время Δt. Будем рассматривать все мéньшие и мéньшие промежутки Δt, устремляя Δt к нулю. Предел

$$\lim_{\Delta t \to 0} \frac{\Delta S(t_0)}{\Delta t} = S'(t_0) = v(t_0)$$

называется *мгновенной скоростью* точки в момент времени t_0.

1.4. Свойства производных и основные теоремы дифференциального исчисления.

Основные свойства производных.

1. Если функция $f(x)$ имеет производную в точке x_0, то и функция $c\cdot f(x)$ (c — постоянная) также имеет производную в точке x_0, причем

$$(c\cdot f(x))'_{x=x_0} = c\cdot f'(x_0),$$

т. е. постоянный множитель можно выносить из-под знака производной.

2. Если функции $f(x)$ и $g(x)$ имеют производные в точке x_0, то и функции $f(x)+g(x)$ и $f(x)-g(x)$ также имеют производные в этой точке, причем

$$(f(x) \pm g(x))'_{x=x_0} = f'(x_0) \pm g'(x_0),$$

т. е. производная суммы (разности) двух функций равна сумме (разности) их производных.

3. Если функции $f(x)$ и $g(x)$ имеют производные в точке x_0, то и функции $f(x)\cdot g(x)$ и $f(x)/g(x)$ также имеют производные в этой

точке, причем

$$(f(x)\cdot g(x))'_{x=x_0}=f'(x_0)\cdot g(x_0)+f(x_0)\cdot g'(x_0),$$

$$\left(\frac{f(x)}{g(x)}\right)'_{x=x_0}=\frac{f'(x_0)\cdot g(x_0)-f(x_0)\cdot g'(x_0)}{g^2(x_0)},\quad g(x_0)\neq 0.$$

4. Если функция $y=f(x)$ имеет производную в точке x_0, а функция $z=g(y)$ имеет производную в точке $y_0=f(x_0)$, то и сложная функция

$$z=g(f(x))$$

имеет производную в точке x_0, причем

$$(g(f(x)))'_{x=x_0}=g'_y(y_0)\cdot f'_x(x_0),$$

где индексы y и x у производных указывают, по какому аргументу вычисляются производные.

5. Пусть функция $y=f(x)$ удовлетворяет условиям существования обратной функции и в точке x_0 имеет конечную и отличную от нуля производную. Тогда обратная функция $x=f^{-1}(y)$ в соответствующей точке $y_0=f(x_0)$ также имеет производную. Производные этих взаимно обратных функций связаны равенством

$$(f^{-1}(y))'_{y=y_0}=\frac{1}{f'(x_0)}.$$

Основные теоремы дифференциального исчисления.

Теорема Ферма. *Пусть функция $f(x)$ определена в некотором промежутке $(a;\ b)$ и во внутренней точке c этого промежутка принимает наибольшее (наименьшее) значение. Если в этой точке существует конечная производная $f'(c)$, то $f'(c)=0$.*

Теорема Ролля. *Если функция $f(x)$ определена и непрерывна на промежутке $[a;\ b]$, имеет конечную производную во всех точках промежутка $(a;\ b)$ и*

$$f(a)=f(b),$$

то в промежутке $(a;\ b)$ найдется такая точка c, что

$$f'(c)=0.$$

Теорема Лагранжа. *Если функция $f(x)$ определена и непрерывна на промежутке $[a;\ b]$ и имеет конечную производную во всех точках промежутка $(a;\ b)$, то найдется такая точка $c\in(a;\ b)$, что*

$$\frac{f(b)-f(a)}{b-a}=f'(c).$$

1.5. Вычисление производных элементарных функций. Производные элементарных функций вычисляются по определению производной (см. формулу (3) п. 1.1), а также с помощью теорем о производных.

Производные показательной и логарифмической функций. Вычислим производную показательной функции $f(x)=a^x$ в точке $x=0$. Составим приращение функции в точке $x=0$:

$$a^{0+\Delta x}-a^0=a^{\Delta x}-1.$$

По определению производной

$$(a^x)'_{x=0}=\lim_{\Delta x\to 0}\frac{a^{\Delta x}-1}{\Delta x}.$$

На основании формулы (3) п. 6.4 гл. 8 получаем

$$\lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x} = \ln a \tag{6}$$

и, следовательно,

$$(a^x)'_{x=0} = \ln a.$$

Производная функции $f(x) = a^x$ во всех остальных точках числовой оси вычисляется следующим образом:

$$(a^x)' = \lim_{\Delta x \to 0} \frac{a^{x+\Delta x} - a^x}{\Delta x} = \lim_{\Delta x \to 0} \frac{a^x \left(a^{\Delta x} - 1 \right)}{\Delta x}.$$

Так как множитель a^x не зависит от Δx, то, вынося его из-под знака предела, на основании формулы (6) получаем

$$\lim_{\Delta x \to 0} \frac{a^x \left(a^{\Delta x} - 1 \right)}{\Delta x} = a^x \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x} = a^x \ln a$$

и, следовательно,

$$(a^x)' = a^x \ln a.$$

Правило вычисления производной показательной функции позволяет дать новое определение числа e (см. п. 2.4 гл. 8).

Числом e называется основание показательной функции, производная которой равна самой функции:

$$(e^x)' = e^x.$$

Теперь по теореме о производной обратной функции (см. свойство 5 п. 1.4) можно вычислить производную логарифмической функции $f(x) = \log_a x$:

$$(\log_a x)' = \frac{1}{x \ln a}.$$

Если $a = e$, то последняя формула принимает следующий простой вид:

$$(\ln x)' = \frac{1}{x}.$$

Производные степенных функций. Используя определение производной, можно доказать, что производная функции $f(x) = c$ (c — постоянная) равна нулю, а производная функции $f(x) = x$ равна единице. Далее, используя теорему о производной произведения функций, методом математической индукции можно доказать, что при любом натуральном n

$$(x^n)' = n \cdot x^{n-1}.$$

Производную степенной функции $f(x) = x^\alpha$ (α — любое действительное число) можно вычислить как производную сложной функции, если записать степенную функцию в виде

$$x^\alpha = e^{\ln \left(x^\alpha \right)} = e^{\alpha \ln x}.$$

По правилу вычисления производной сложной функции

$$\left(x^\alpha \right)' = \left(e^{\alpha \ln x} \right)' = e^{\alpha \ln x} (\alpha \ln x)' = e^{\alpha \ln x} \cdot \alpha \cdot \frac{1}{x} = x^\alpha \cdot \alpha \frac{1}{x} = \alpha x^{\alpha - 1}.$$

В качестве более сложного примера на использование теоремы о производной сложной функции вычислим производную «показательно-степенной» функции $f(x) = x^x$. Записывая функцию в виде $x^x = e^{\ln (x^x)} = e^{x \ln x}$, имеем

$$(x^x)' = \left(e^{x \ln x} \right)' = e^{x \ln x} (x \cdot \ln x)' =$$

$$= e^{x \ln x} \left(\ln x + x \cdot \frac{1}{x} \right) = e^{x \ln x} (\ln x + 1) = x^x (\ln x + 1).$$

Производные тригонометрических функций. Вычислим производную функции $f(x)=\sin x$. Пусть x_0 — произвольная точка числовой оси. Составим приращение функции

$$\Delta \sin x_0=\sin (x_0+\Delta x)-\sin x_0.$$

Разность, стоящую в правой части равенства, запишем в виде произведения

$$\sin (x_0+\Delta x)-\sin x_0=2\cos\left(x_0+\frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2}.$$

По определению производной

$$\left(\sin x\right)'_{x=x_0}=\lim_{\Delta x\to 0}\frac{\Delta \sin x_0}{\Delta x}=\lim_{\Delta x\to 0}\frac{2\cos\left(x_0+\frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2}}{\Delta x}=$$

$$=\lim_{\Delta x\to 0}\cos\left(x_0+\frac{\Delta x}{2}\right)\cdot\lim_{\Delta x\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}.$$

Первый предел, стоящий в правой части последнего равенства, вычислим, используя свойство непрерывности функции $\cos x$:

$$\lim_{\Delta x\to 0}\cos\left(x_0+\frac{\Delta x}{2}\right)=\cos x_0,$$

а второй — с помощью предела 1) п. 6.4 гл. 8, на основании которого получаем

$$\lim_{\Delta x\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}=1.$$

Таким образом,

$$\left(\sin x\right)'_{x=x_0}=\cos x_0.$$

Так как точка x_0 — произвольная точка числовой оси, то

$$(\sin x)'=\cos x. \tag{7}$$

С помощью формулы (7) и теорем о производных можно вычислить производные остальных тригонометрических функций. Например, производная функции $f(x)=\cos x$ вычисляется следующим образом: представляя функцию $\cos x$ в виде

$$\cos x=\sin\left(\frac{\pi}{2}-x\right),$$

вычислим ее производную с помощью теоремы о производной сложной функции:

$$(\cos x)'=\left(\sin\left(\frac{\pi}{2}-x\right)\right)'=\cos\left(\frac{\pi}{2}-x\right)\cdot\left(\frac{\pi}{2}-x\right)'=-\cos\left(\frac{\pi}{2}-x\right)=-\sin x.$$

Производные функций $\operatorname{tg} x$ и $\operatorname{ctg} x$ вычисляются с помощью теоремы о производной частного.

Производные обратных тригонометрических функций.

Производные обратных тригонометрических функций $f(x)=\arcsin x$, $f(x)=\arccos x$, $f(x)=\operatorname{arctg} x$ и $f(x)=\operatorname{arcctg} x$ вычисляются с помощью теоремы о производной обратной функции п. 1.4. Например, производная функции $f(x)=\arcsin x$ вычисляется следующим образом: обратной функцией для функции $y=\arcsin x$ является функция $x=\sin y$; по теореме о производной обратной функции имеем

$$(\arcsin x)'=\frac{1}{(\sin y)'}=\frac{1}{\cos y}=\frac{1}{\cos(\arcsin x)}.$$

Так как при любом значении x, принадлежащем области определения функции $\arcsin x$, $\cos(\arcsin x)=\sqrt{1-x^2}$, то

$$(\arcsin x)'=\frac{1}{\sqrt{1-x^2}}.$$

Производные элементарных функций приведены в табл. 8.

Таблица 8

Функция	Производная
$f(x)=c$	$f'(x)=0$
$f(x)=x^\alpha$	$f'(x)=\alpha x^{\alpha-1}$
$f(x)=a^x$	$f'(x)=a^x\ln a$
$f(x)=e^x$	$f'(x)=e^x$
$f(x)=\log_a x$	$f'(x)=\frac{1}{x\cdot\ln a}$
$f(x)=\ln x$	$f'(x)=\frac{1}{x}$
$f(x)=\sin x$	$f'(x)=\cos x$
$f(x)=\cos x$	$f'(x)=-\sin x$
$f(x)=\operatorname{tg} x$	$f'(x)=\frac{1}{\cos^2 x}$
$f(x)=\operatorname{ctg} x$	$f'(x)=-\frac{1}{\sin^2 x}$
$f(x)=\arcsin x$	$f'(x)=\frac{1}{\sqrt{1-x^2}}$
$f(x)=\arccos x$	$f'(x)=-\frac{1}{\sqrt{1-x^2}}$
$f(x)=\operatorname{arctg} x$	$f'(x)=\frac{1}{1+x^2}$
$f(x)=\operatorname{arcctg} x$	$f'(x)=-\frac{1}{1+x^2}$

1.6. Производные высшего порядка. Пусть функция $f(x)$ дифференцируема в промежутке $(a;\ b)$. Производная данной функции $f'(x)$ представляет собой функцию аргумента x. Выберем некоторую точку $x_0\in(a;\ b)$ и составим приращение аргумента $\Delta x=x-x_0$ и приращение функции $f'(x)$ в точке x_0:

$$\Delta f'(x_0)=f'(x_0+\Delta x)-f'(x_0).$$

Рассмотрим предел

$$\lim_{\Delta x\to 0}\frac{f'(x_0+\Delta x)-f'(x_0)}{\Delta x}.$$

Если этот предел существует, то говорят, что функция $f(x)$ имеет *производную второго порядка* (или *вторую производную*) в точке x_0. Производную второго порядка функции $f(x)$ в точке x_0 обозначают

$$f''(x_0),\quad \text{или}\quad \frac{d^2f(x_0)}{dx^2},\quad \text{или}\quad f''(x)\big|_{x=x_0}.$$

Таким образом,

$$f''(x_0) = \lim_{\Delta x \to 0} \frac{f'(x_0 + \Delta x) - f'(x_0)}{\Delta x}.$$

Производная второго порядка, так же как и производная первого порядка, допускает простое кинематическое истолкование.

Пусть материальная точка движется под действием сил по некоторой траектории (рис. 9.3), и пусть в момент времени t_0 она занимает положение M_0, а в момент времени $t_1 = t_0 + \Delta t$ — положение M_1. Путь, пройденный точкой от положения M_0 до положения M_1, равен длине $\Delta S(t_0)$ части траектории M_0M_1. Скорость точки в каждый момент времени является производной пройденного пути:

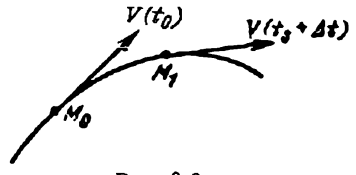


Рис. 9.3

$$v(t_0) = S'(t_0), \quad v(t_1) = S'(t_1),$$

направленной по касательной к траектории. Предел

$$\lim_{\Delta t \to 0} \frac{v(t_1) - v(t_0)}{\Delta t} = \lim_{\Delta t \to 0} \frac{v(t_0 + \Delta t) - v(t_0)}{\Delta t} =$$

$$= \lim_{\Delta t \to 0} \frac{S'(t_0 + \Delta t) - S'(t_0)}{\Delta t} = a(t_0)$$

называется *мгновенным ускорением* точки в момент времени t_0 и представляет собой вторую производную от пути $S(t)$:

$$a(t) = S''(t).$$

Далее, если для функции $f''(x)$ в некоторой точке x существует предел

$$\lim_{\Delta x \to 0} \frac{f''(x_0 + \Delta x) - f''(x_0)}{\Delta x},$$

то говорят, что функция $f(x)$ имеет *производную третьего порядка* (или *третью производную*) в точке x_0:

$$f'''(x_0) = \lim_{\Delta x \to 0} \frac{f''(x_0 + \Delta x) - f''(x_0)}{\Delta x}.$$

Аналогичным способом определяются производные четвертого, пятого и т. д. порядков. Производная *n-го порядка* функции $f(x)$ обычно обозначается

$$f^{(n)}(x) \quad \text{или} \quad \frac{d^n f}{dx^n}.$$

Производные n-го порядка суммы и произведения двух функций $f(x)$ и $g(x)$ вычисляются по формулам

$$(f \pm g)^{(n)} = f^{(n)} \pm g^{(n)},$$

$$(f \cdot g)^{(n)} = C_n^0 f^{(n)} \cdot g^{(0)} + C_n^1 \cdot f^{(n-1)} g + C_n^2 f^{(n-2)} g'' + \ldots$$

$$\ldots + C_n^{n-1} f' g^{(n-1)} + C_n^n f^{(0)} g^{(n)} = \sum_{i=0}^{n} C_n^i f^{(n-i)} g^{(i)}.$$

В последней формуле (называемой *формулой Лейбница*) через $f^{(0)}$ и $g^{(0)}$ обозначены функции $f(x)$ и $g(x)$ соответственно, C_n^i — число сочетаний из n элементов по i элементов (биномиальные коэффициенты).

1.7. Производные функции, заданной в параметрической форме.

Первая производная. Пусть функция задана в параметрическом виде

$$x=\varphi(t), \quad y=\psi(t) \qquad (t-\text{параметр}) \tag{8}$$

и $\varphi(t)$ и $\psi(t)$ — дифференцируемые функции на промежутке $(a; b)$. Если в некоторой точке $t_0 \in (a; b)$ производная функции $x=\varphi(t)$ отлична от нуля, то в достаточно малой окрестности точки $x_0=\varphi(t_0)$ параметрически заданная функция (8) может быть задана в явной форме

$$y=\psi(t(x))=y(x), \tag{9}$$

где $t=t(x)$ — обратная функции $x=\varphi(t)$.

Дифференцируя функцию (9) как сложную функцию x и используя связь между производными двух взаимно обратных функций, получаем формулу для вычисления производной параметрически заданной функции

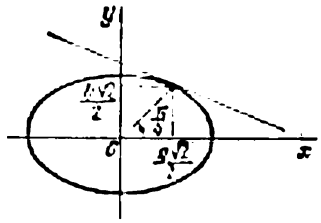


Рис. 9.4

$$y'_x=\psi'_t\, t'_x=\frac{\psi'_t}{\varphi'_t} \tag{10}$$

или

$$\frac{dy}{dx}=\frac{dy}{dt}\Big/\frac{dx}{dt}.$$

Полученное выражение для первой производной функции, заданной в параметрическом виде, позволяет вычислить тангенс угла наклона касательной к графику функции

$$\operatorname{tg}\alpha=\frac{y'_t}{x'_t}, \tag{11}$$

а также написать уравнение касательной к графику параметрически заданной функции $x=\varphi(t)$, $y=\psi(t)$.

Пример. Написать уравнение касательной к графику функции, заданной в параметрической форме:

$$x=a\sin t, \quad y=b\sin t, \quad t\in[0,\ 2\pi), \quad a>0, \quad b>0, \tag{12}$$

при $t=\frac{\pi}{4}$.

Нетрудно проверить, что график данной функции представляет собой кривую второго порядка — эллипс (рис. 9.4), так как из формул (12) следует, что

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1.$$

Тангенс угла наклона касательной в соответствии с формулой (11) равен

$$\operatorname{tg}\alpha=\frac{y'_t}{x'_t}\Big|_{t=\frac{\pi}{4}}=-\frac{b\cos t}{a\sin t}\Big|_{t=\frac{\pi}{4}}=-\frac{b}{a}.$$

По условию задачи касательная проходит через точку с координатами $x_0=a\cos\frac{\pi}{4}=\frac{a\sqrt{2}}{2}$, $y_0=b\sin\frac{\pi}{4}=\frac{b\sqrt{2}}{2}$, и, следовательно, уравнение касательной имеет вид

$$y-\frac{b\sqrt{2}}{2}=-\frac{b}{a}\left(x-\frac{a\sqrt{2}}{2}\right).$$

Вторая производная. Для вычисления второй производной параметрически заданной функции воспользуемся полученной выше формулой (10)

$$y''_{xx}=(y'_x)'_x=(\psi'_t\cdot t'_x)'_x=(\psi'_t)'_x\cdot t'_x+\psi'_t\cdot(t'_x)'_x.$$

Вычисляя производную по x от функции ψ'_t как производную сложной функции

$$(\psi'_t)'_x=\psi''_{tt}\cdot t'_x,$$

получаем

$$y''_{xx}=\psi''_{tt}\cdot(t'_x)^2+\psi'_t t''_{xx}.$$

Так как $t'_x=1/x'_t$, а

$$t''_{xx}=(t'_x)'_x=\left(\frac{1}{x'_t}\right)'_x=-\frac{1}{(x'_t)^2}\cdot x''_{tt}\cdot t'_x=-\frac{1}{(x'_t)^3}\cdot x''_{tt},$$

окончательно получаем

$$y''_{xx}=\frac{y''_{tt}x'_t-y'_t x''_{tt}}{(x'_t)^3}.$$

1.8. Понятие частной производной. Пусть $z=f(x, y)$ — некоторая функция двух переменных, определенная в открытой области D, и пусть точка $M(x_0; y_0)\in D$. Положим, что значение аргумента y функции $f(x, y)$ равно y_0. Тогда функцию $f(x, y_0)$ можно рассматривать как функцию одной переменной x. Обозначим через Δx приращение аргумента x:

$$\Delta x=x-x_0.$$

Приращение функции $f(x, y)$ при $y=y_0$ обозначим

$$\Delta_x z=f(x_0+\Delta x, y_0)-f(x_0, y_0).$$

Приращение $\Delta_x z$ называют *частным приращением* функции $f(x, y)$ по x. Составим отношение

$$\frac{f(x_0+\Delta x, y_0)-f(x_0, y_0)}{\Delta x},$$

которое при фиксированных x_0 и y_0 является функцией приращения Δx.

Если существует предел

$$\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x, y_0)-f(x_0, y_0)}{\Delta x},$$

то говорят, что функция $f(x, y)$ имеет *частную производную* по x в точке $(x_0; y_0)$, и пишут

$$\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x, y_0)-f(x_0, y_0)}{\Delta x}=f'_x(x_0, y_0).$$

Частную производную $f'_x(x_0, y_0)$ также обозначают

$$z'_x, \quad \text{или} \quad \frac{\partial z}{\partial x}, \quad \text{или} \quad \frac{\partial f(x_0, y_0)}{\partial x}.$$

Аналогично, полагая $x=x_0$, можно рассмотреть предел

$$\lim_{\Delta y\to 0}\frac{f(x_0, y_0+\Delta y)-f(x_0, y_0)}{\Delta y}.$$

Если этот предел существует, то говорят, что функция $f(x, y)$ имеет *частную производную* по y в точке $(x_0; y_0)$.

Эту частную производную обозначают

$$f'_y(x_0, y_0), \quad \text{или} \quad z'_y, \quad \text{или} \quad \frac{\partial z}{\partial y}, \quad \text{или} \quad \frac{\partial f(x_0, y_0)}{\partial y}.$$

Говорят, что функция имеет частную производную в открытой области D, если она имеет частную производную в каждой точке этой области.

Вычисление частной производной функции нескольких переменных производится по тем же правилам, что и вычисление обыкновенной производной.

Пример 1. Пусть $z=x^y$ $(x>0)$. Частная производная по x этой функции вычисляется при $y=\text{const}$, т. е. представляет собой производную степенной функции:

$$\frac{\partial z}{\partial x}=y\cdot x^{y-1}.$$

Частная производная по y вычисляется при $x=\text{const}$, т. е. вычисляется как обыкновенная производная показательной функции:

$$\frac{\partial z}{\partial y}=x^y\ln x.$$

Пример 2. Частные производные функции

$$z=\frac{x^2+x+1}{\ln y}$$

равны

$$\frac{\partial z}{\partial x}=\frac{1}{\ln y}(x^2+x+1)'=\frac{2x+1}{\ln y};$$

$$\frac{\partial z}{\partial y}=(x^2+x+1)\cdot\left(\frac{1}{\ln y}\right)'=(x^2+x+1)\left(-\frac{1}{\ln^2 y}\right)\cdot(\ln y)'=-\frac{x^2+x+1}{y\ln^2 y},$$

т. е. вычисляются, как обыкновенные производные по переменной x (в предположении, что y — постоянная) и по переменной y (в предположении, что x — постоянная).

Если функция $z=f(x, y)$ имеет во всех точках некоторой открытой области D частную производную по одной из переменных, то эта производная, сама являясь функцией переменных x и y, может в свою очередь в некоторой точке $(x_0; y_0)\in D$ иметь частные производные по x или по y. Для данной функции $z=f(x, y)$ эти производные будут частными производными второго порядка.

Так, например, рассмотрим предел

$$\lim_{\Delta x\to 0}\frac{f'_x(x_0+\Delta x, y_0)-f'_x(x_0, y_0)}{\Delta x}.$$

Если этот предел существует, то говорят, что функция $f(x, y)$ имеет *вторую частную производную по x* (или частную производную *второго порядка*) в точке (x_0, y_0). Эту производную обозначают $f''_{xx}(x_0, y_0)$

$$\lim_{\Delta x\to 0}\frac{f'_x(x_0+\Delta x, y_0)-f'_x(x_0, y_0)}{\Delta x}=f''_{xx}(x_0, y_0).$$

Частную производную второго порядка по x функции $z=f(x, y)$ также обозначают

$$\frac{\partial^2 f}{\partial x^2},\quad \text{или}\quad \frac{\partial^2 z}{\partial x^2},\quad \text{или}\quad z''_{xx}.$$

Аналогично определяются производные

$$f''_{yy},\quad f''_{xy},\quad f''_{yx}.$$

Производные f''_{xy} и f''_{yx} называют *смешанными частными производными*.

Условие равенства смешанных частных производных. Если в некоторой точке $(x_0; y_0)$, принадлежащей открытой области D, смешанные частные производные f''_{xy} и f''_{yx} существуют и непрерывны как функции переменных x, y, то они равны:

$$f''_{xy}(x_0, y_0)=f''_{yx}(x_0, y_0).$$

§ 2. Дифференциал

2.1. Понятие дифференциала. Пусть $y=f(x)$ — некоторая функция, определенная на промежутке $(a; b)$ и непрерывная в некоторой точке x_0 этого промежутка, и пусть приращению аргумента Δx соответствует приращение функции

$$\Delta y=\Delta f(x_0)=f(x_0+\Delta x)-f(x_0),$$

которое можно рассматривать как функцию аргумента Δx.

Если для приращения функции Δy существует такое число $A\neq 0$, что приращение функции может быть представлено в виде

$$\Delta y=A\cdot\Delta x+\omega(\Delta x),$$

где слагаемое $\omega(\Delta x)$ удовлетворяет равенству

$$\lim_{\Delta x\to 0}\frac{\omega(\Delta x)}{\Delta x}=0,$$

то говорят, что функция $f(x)$ *дифференцируема* в точке x_0.

Выражение $A\cdot\Delta x$ называется *дифференциалом функции* $y=f(x)$ в точке x_0 и обозначается

$$dy \quad \text{или} \quad df(x_0).$$

Иногда дают следующее краткое определение дифференциала. *Дифференциал — это главная линейная часть приращения функции* *).

Связь между дифференцируемостью функции и существованием ее производной устанавливается следующей теоремой.

Для того чтобы функция $y=f(x)$ в точке x_0 была дифференцируема, необходимо и достаточно, чтобы она имела в этой точке конечную производную $f'(x_0)$. При выполнении этого условия значение производной $f'(x_0)$ равно числу A и дифференциал dy функции $y=f(x)$ равен

$$dy=f'(x_0)\,\Delta x. \tag{1}$$

В равенстве (1) Δx не обязательно бесконечно малая; но если Δx бесконечно малая, то и dy бесконечно малая и именно в этих случаях dy (при условии, что $f'\neq 0$) является главной частью бесконечно малого приращения функции Δy.

Дифференциалом независимой переменной x в точке x_0 называют ее приращение Δx, т. е.

$$dx=\Delta x.$$

С учетом этого соглашения формула (1) может быть переписана в виде

$$dy=f'_x\cdot dx,$$

или

$$f'_x=\frac{dy}{dx}.$$

Выражение, стоящее в правой части последнего равенства, которое раньше (см. п. 1.1) было просто обозначением для производной, теперь можно понимать как дробь с числителем dy и знаменателем dx.

*) Определение понятия главной части бесконечно малой величины см. в п. 7.2 гл. 8.

Правила вычисления дифференциалов суммы, произведения и частного двух функций аналогичны соответствующим правилам вычисления производных:

$$d(c\cdot f(x)) = c\,df(x) \qquad (c \text{ постоянная});$$
$$d(f(x) \pm g(x)) = df(x) \pm dg(x);$$
$$d(f(x)\cdot g(x)) = g(x)\,df(x) + f(x)\cdot dg(x);$$
$$d\left(\frac{f(x)}{g(x)}\right) = \frac{g(x)\,df(x) - f(x)\,dg(x)}{g^2(x)}.$$

2.2. Геометрическая интерпретация дифференциала. Пусть $y = f(x)$ — некоторая функция, M — точка, принадлежащая графику функции и с координатами $(x_0;\ f(x_0))$. Проведем через точку M (рис. 9.5) касательную к графику функции. Угловой коэффициент наклона касательной (тангенс угла наклона) равен значению производной $f'(x_0)$. Если аргументу функции придать приращение Δx, то приращение функции Δy будет равно

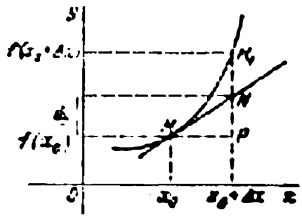


Рис. 9.5

$$\Delta y = f(x_0 + \Delta x) - f(x_0).$$

На рис. 9.5 приращение функции Δy — длина отрезка M_1P.

При том же самом приращении аргумента Δx приращение касательной будет равно длине отрезка NP. Вычисляя $|NP|$ как катет прямоугольного треугольника MNP, находим

$$|NP| = |MP|\cdot \operatorname{tg}\alpha = f'(x_0)\cdot\Delta x.$$

По определению дифференциала $f'(x_0)\cdot\Delta x = dy$, следовательно, длина отрезка $|NP|$ и есть дифференциал функции dy.

Таким образом, в то время как Δy есть приращение ординаты графика функции, dy является приращением ординаты касательной.

2.3. Инвариантность дифференциала. Пусть $y = f(x)$ и $x = g(t)$ — некоторые функции указанных аргументов такие, что из них может быть образована сложная функция

$$y = f(g(t)).$$

Производная сложной функции $y(t)$ вычисляется по формуле

$$y'_t = f'_x\cdot g'_t,$$

где индексы указывают, по какому аргументу вычисляются производные.

Дифференциал функции $y = f(x)$ в предположении, что x — независимая переменная, определяется равенством

$$dy = f'_x\,dx. \tag{2}$$

Если рассматривать функцию y как функцию независимой переменной t, то ее дифференциал определяется равенством

$$dy = y'_t\,dt.$$

Подставляя в это равенство $y'_t = f'_x\cdot g'_t$, получаем

$$dy = f'_x\cdot g'_t\,dt. \tag{3}$$

Но, с другой стороны, $g'_t\,dt = dg = dx$; следовательно, равенство (3) записывается в виде

$$dy = f'_x\,dx.$$

Таким образом, предполагая, что y есть функция аргумента t, мы получаем то же самое выражение (2) для дифференциала dy, т. е. форма дифференциала не зависит от того, является ли x независимой переменной, или она в свою очередь является функцией какой-либо переменной.

Это свойство называется свойством *инвариантности формы дифференциала.*

Пример. Пусть $y=\sin x$ и $x=\ln t$ — две функции, из которых можно образовать сложную функцию

$$y=\sin(\ln t).$$

Производная этой функции равна

$$y'_t=\cos(\ln t)\cdot(\ln t)'=\frac{\cos(\ln t)}{t},$$

а дифференциал

$$dy=y_t'dt=\frac{\cos(\ln t)}{t}\,dt.$$

С другой стороны,

$$dx=d(\ln t)=\frac{1}{t}\,dt$$

Преобразуя с помощью этого равенства правую часть предыдущей формулы, получаем также следующее выражение для дифференциала dy:

$$dy=y_t'\,dt=\cos(\ln t)\,d(\ln t),$$

или

$$dy=\cos x\,dx.$$

Но точно такое же выражение для дифференциала dy получается и в том случае, когда в первой из исходных функций ($y=\sin x$) x считается независимой переменной.

§ 3. Исследование функций. Построение графика функции

3.1. Условия постоянства и монотонности функций. Пусть функция $f(x)$ определена и дифференцируема в промежутке $(a;\ b)$. Для того чтобы функция в промежутке $(a;\ b)$ была постоянна, необходимо и достаточно, чтобы ее производная была равна нулю во всем промежутке $(a;\ b)$:

$$f'(x)=0.$$

Пусть функция $f(x)$ определена и дифференцируема в промежутке $(a;\ b)$. Для того чтобы функция была возрастающей в промежутке $(a;\ b)$, достаточно, чтобы

$$f'(x)>0 \quad \text{для всех} \quad x\in(a;\ b).$$

Для убывания функции достаточно, чтобы

$$f'(x)<0 \quad \text{для всех} \quad x\in(a;\ b).$$

Сформулированные условия не являются необходимыми условиями возрастания и убывания функции.

Например, функция $f(x)=x^3$ — возрастающая, однако в точке $x=0$ ее производная равна нулю.

3.2. Максимумы и минимумы функции. Говорят, что функция $f(x)$ имеет в точке x_0 *максимум* (или *минимум*), если найдется такая δ-окрестность точки x_0, принадлежащая области определения функции, что для всех $x \neq x_0$, принадлежащих промежутку $(x_0-\delta;\ x_0+\delta)$, выполняется неравенство

$$f(x) < f(x_0) \qquad (\text{соответственно } f(x) > f(x_0)).$$

На рис. 9.6 изображена функция, определенная на промежутке $[a;\ b]$, имеющая в точках x_1 и x_3 максимумы, а в точках x_2 и x_4 — минимумы.

Точки максимума и минимума называются точками *экстремума* функции, а значения функции в этих точках называются *экстремальными значениями функции*. Для функции, график которой изображен на рис. 9.6, точки a и b не являются точками экстремума функции, так как в любой сколь угодно малой окрестности этих точек найдутся точки, не принадлежащие области определения $f(x)$.

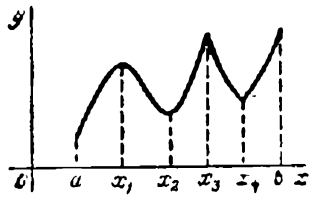


Рис. 9.6

Необходимое условие существования экстремума функции. Пусть функция $f(x)$ дифференцируема в промежутке $(a;\ b)$. Если в некоторой точке $x_0 \in (a;\ b)$ функция $f(x)$ имеет экстремум, то в этой точке производная равна нулю *):

$$f'(x_0) = 0.$$

Обратное утверждение, вообще говоря, неверно — если в точке x_0 производная функции $f(x)$ равна нулю, то функция $f(x)$ в точке x_0 может не иметь экстремума. Так, производная функции $f(x) = x^3$ в точке $x_0 = 0$ равна нулю, но функция не имеет экстремума в этой точке.

Достаточное условие существования экстремума функции. Пусть функция $f(x)$ определена и непрерывна в промежутке $(a;\ b)$ и на всем промежутке $(a;\ b)$ (за исключением, быть может, конечного числа точек) дифференцируема.

Критическими точками функции в промежутке $(a;\ b)$ называются точки, в которых производная функции либо равна нулю, либо не существует.

Пусть точка $x_0 \in (a;\ b)$ — критическая точка функции $f(x)$ и в некоторой δ-окрестности этой точки, по крайней мере для $x \neq x_0$, существует конечная производная $f'(x)$, которая слева и справа от x_0 сохраняет знак (возможно, с разных сторон — разный). Тогда возможны следующие три случая:

2) $f'(x) > 0$ при $x < x_0$ и $f'(x) < 0$ при $x > x_0$, т. е. производная $f'(x)$ при переходе через точку x_0 меняет знак с плюса на минус. В этом случае в промежутке $(x_0-\delta,\ x_0)$ функция $f(x)$ возрастает, а в промежутке $(x_0;\ x_0+\delta)$ убывает, так что значение $f(x_0)$ будет наибольшим в промежутке $(x_0-\delta;\ x_0+\delta)$, т. е. в точке x_0 функция имеет максимум;

*) Если расширить класс рассматриваемых функций, полагая, что в конечном числе точек промежутка $(a;\ b)$ функция $f(x)$ не имеет конечной производной, то экстремум функции может приходиться на точку, в которой функция не имеет конечной производной.

2) $f'(x) < 0$ при $x < x_0$ и $f'(x) > 0$ при $x > x_0$, т. е. производная $f'(x)$ при переходе через точку x_0 меняет знак с минуса на плюс. В этом случае в точке x_0 функция имеет минимум;

3) $f'(x) > 0$ как при $x > x_0$, так и при $x < x_0$, либо $f'(x) < 0$ и слева и справа от x_0, т. е. при переходе через точку x_0 производная $f'(x)$ не меняет знака. Тогда функция либо возрастает, либо убывает на промежутке $(x_0 - \delta; x_0 + \delta)$; в любой сколь угодно малой окрестности точки x_0 с одной стороны найдутся точки x, в которых

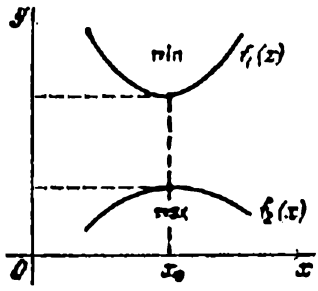


Рис. 9.7

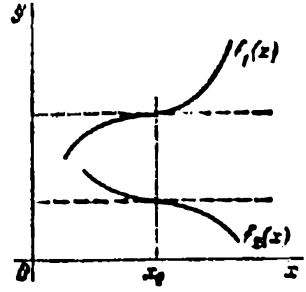


Рис. 9.8

$f(x) < f(x_0)$, а с другой — точки x, в которых $f(x) > f(x_0)$, и в точке x_0 экстремума нет.

Первое правило для проверки критической точки x_0 на экстремум заключается в том, что, подставляя в производную $f'(x)$ сначала $x < x_0$, а затем $x > x_0$, устанавливают знак производной в окрестности точки x_0 слева и справа от нее. Если при этом производная $f'(x)$ меняет знак с плюса на минус, то в точке x_0 — максимум,

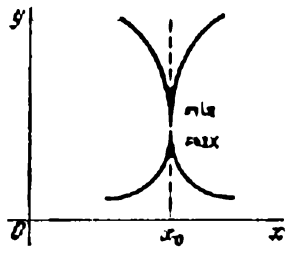


Рис. 9.9

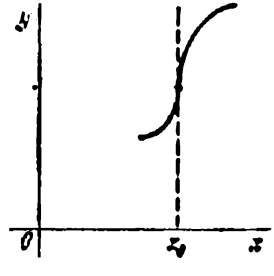


Рис. 9.10

если с минуса на плюс — то минимум; если же знак не меняется, то экстремума нет.

На рис. 9.7 производная функции $f_1(x)$ при переходе через точку x_0 меняет знак с минуса на плюс и функция $f_1(x)$ в точке x_0 имеет минимум, а функция $f_2(x)$, производная которой при переходе через точку x_0 меняет знак с плюса на минус,— максимум.

На рис. 9.8 как функция $f_1(x)$, так и функция $f_2(x)$ имеют в точке x_0 производную, равную нулю: при переходе через точку x_0 производные той и другой функции знака не меняют и экстремумов у функций $f_1(x)$ и $f_2(x)$ в точке x_0 нет.

На рис. 9.9 и 9.10 изображены графики функций, которые не имеют производных в точке x_0. В согласии с правилами смены знака

производных для этих функций, они имеют соответственно минимум, максимум и не имеют экстремумов в точке x_0.

Примером функции, не имеющей производной в точке и имеющей в этой точке минимум, является функция $y=|x|$ (см. п. 1.1).

Еще раз подчеркнем, что сформулированное выше достаточное условие существования экстремума функции в точке x_0 пригодно как в случаях, когда производная в точке x_0 обращается в нуль, так и в случаях, когда в этой точке производная не существует, и применимо к непрерывным функциям, имеющим непрерывную производную $f'(x)$ в промежутке $(a; b)$, за исключением, быть может, конечного числа точек.

Сформулируем еще одно достаточное условие существования экстремума функций в точке. Точку x_0 называют *стационарной* точкой функции $f(x)$, если в этой точке функция $f(x)$ дифференцируема и ее производная равна нулю.

Пусть x_0 — стационарная точка функции $f(x)$ и в некоторой окрестности точки x_0, включая саму эту точку, функция $f(x)$ имеет производную, а в точке x_0 существует и вторая производная $f''(x_0)$. Если вторая производная f'' в точке x_0 положительна, то функция $f(x)$ в точке x_0 имеет минимум; если вторая производная f'' в точке x_0 отрицательна,— то максимум.

Сформулированное правило имеет, вообще говоря, более узкий круг применения по сравнению с первым правилом; оно, например, явно неприменимо к точкам, в которых не существует первая производная. В тех случаях, когда вторая производная в точке x_0 обращается в нуль, правило также ничего не дает и решение вопроса о существовании в этой точке экстремума зависит от поведения высших производных.

3.3. Наибольшее и наименьшее значения функции. Пусть функция $f(x)$ определена и непрерывна на конечном замкнутом промежутке $[a; b]$.

Для нахождения наибольшего (наименьшего) значения функции необходимо найти все максимальные (минимальные) значения функции в промежутке $(a; b)$, выбрать из них наибольшее (наименьшее) и сравнить его со значениями функции в точках a и b. Наибольшее (наименьшее) из этих чисел и будет наибольшим (наименьшим) значением функции $f(x)$ на промежутке $[a; b]$. При нахождении наибольшего или наименьшего значения функции может оказаться, что на промежутке $(a; b)$ функция не имеет критических точек. Это говорит о том, что в рассматриваемом промежутке функция возрастает или убывает и, следовательно, достигает своего наибольшего и наименьшего значения на концах промежутка.

Примеры. 1. Найти наибольшее и наименьшее значения функции

$$f(x)=\frac{4}{3}x^3-4x$$

на промежутке $[0; 2]$.

Вычислим производную данной функции

$$f'(x)=4x^2-4.$$

При всех $x\in(0; 2)$ производная существует и равна нулю в точках $x=1$ и $x=-1$. В промежутке $(0; 2)$ лежит лишь одна из этих точек: $x=1$. Значение функции в этой точке равно $-8/3$. При переходе через точку $x=1$ производная меняет знак с минуса на плюс, а потому в точке $x=1$ — минимум функции $f(x)$. Значения функции на концах промежутка: $f(0)=0$, $f(2)=8/3$. Следовательно, наибольшее

значение данной функции $f(x)$ достигается на правом конце промежутка в точке $x=2$, а наименьшее — внутри промежутка в точке $x=1$.

2. Из квадратного листа со стороной a, вырезая по углам равные квадраты и сгибая края (рис. 9.11), составляют прямоугольную открытую коробку. Чему должна быть равна сторона вырезаемого квадрата, чтобы получившаяся коробка имела наибольший объем?

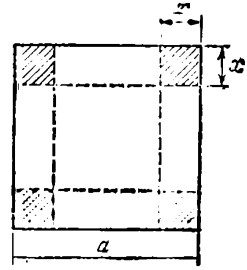


Рис. 9.11

Обозначим сторону вырезаемого квадрата через x. Тогда объем коробки V будет равен

$$x(a-2x)^2,$$

причем x принадлежит промежутку $[0;\ a/2]$. Производная данной функции $V'(x)=(a-2x)(a-6x)$ на промежутке $[0;\ a/2)$ обращается в нуль в точке $x=a/6$. Вычислим вторую производную: $V''(x)=24x-8a$. В точке $x=a/6$ вторая производная отрицательна, и, следовательно, в точке $x=a/6$ функция $V(x)$ имеет максимум. На концах промежутка $[0;\ a/2]$ функция $V(x)$ обращается в нуль. Таким образом, объем коробки при $x=a/6$ будет наибольшим.

3.4. Направление вогнутости кривой. Пусть функция $f(x)$ определена и непрерывна на промежутке $(a;\ b)$, а в точке $x_0 \in (a;\ b)$ имеет конечную производную. Тогда к графику данной функции в точке $M_0(x_0;\ f(x_0))$ можно провести касательную, уравнение которой имеет вид

$$y-y_0=f'(x_0)(x-x_0).$$

Говорят, что в точке M_0 кривая (график функции $y=f(x)$) направлена *вогнутостью* в определенную сторону (*вверх* или *вниз*) от касательной, если в достаточно малой окрестности точки M_0 все точки кривой лежат по одну сторону от касательной (рис. 9.12 и 9.13).

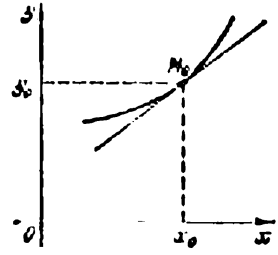


Рис. 9.12

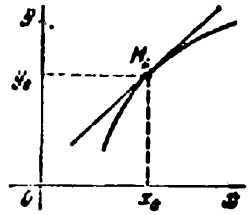


Рис. 9.13

Точку M_0 называют *точкой перегиба*, если в достаточно малой ее окрестности точки кривой с абсциссами $x < x_0$ лежат по одну сторону от касательной, а точки с абсциссами $x > x_0$ — по другую, т. е. в точке M_0 кривая переходит с одной стороны касательной на другую (рис. 9.14).

Направление вогнутости кривой $y=f(x)$ в некоторой точке x_0 выясняется по следующему правилу:

1) если в некоторой δ-окрестности точки x_0 вторая производная $f''(x)$ сохраняет знак плюс (как слева, так и справа от точки x_0), то кривая в точке x_0 направлена вогнутостью вверх (рис. 9.12);

2) если в некоторой δ-окрестности точки x_0 вторая производная сохраняет знак минус (как слева, так и справа от точки x_0), то кривая в точке x_0 направлена вогнутостью вниз (рис. 9.13);

3) если вторая производная меняет знак при переходе через точку x_0, то в точке x_0 график функции $f(x_0)$ имеет перегиб (рис. 9.14).

Таким образом, если ограничиться достаточно малой δ-окрестностью точки x_0, точка перегиба M_0 как бы отделяет те точки, где

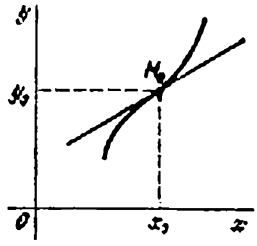


Рис. 9.14

Рис. 9.15

вогнутость кривой направлена вверх, от точек с направлением вогнутости вниз.

Сформулированное правило — достаточное (но не необходимое) условие перегиба кривой. Так, функция $f(x)=x^4$ в точке $x=0$ имеет вторую производную, равную нулю, но тем не менее в этой точке (и вообще при всех $x \in \mathbb{R}$) кривая направлена вогнутостью вверх (рис. 9.15).

Пример. Рассмотрим функцию $f(x)=\sin x$. Вторая производная этой функции равна

$$f''(x)=-\sin x$$

В промежутках $(2\pi k;\ \pi+2\pi k)$ функция $\sin x$ сохраняет знак плюс, $f''(x)=-\sin x<0$ и синусоида направлена вогнутостью вниз. В промежутках $(\pi+2\pi k;\ 2\pi+2\pi k)$ $\sin x<0$ и, следовательно, синусоида направлена вогнутостью вверх. При

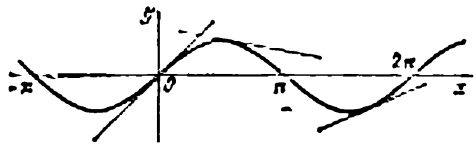


Рис. 9.16

$x=\pi k$ вторая производная обращается в нуль, меняя при этом знак; в этих точках график функции $y=\sin x$ имеет перегиб (рис. 9.16).

3.5. Асимптоты графика функции. Пусть $y=f(x)$ — функция, график которой имеет бесконечную ветвь, т. е. ветвь, имеющую точки, принадлежащие графику функции и находящиеся сколь угодно далеко от начала координат.

Асимптотой графика функции $y=f(x)$ называют прямую, обладающую тем свойством, что расстояние от точки $(x;\ f(x))$ до этой прямой стремится к нулю при движении этой точки вдоль ветви к бесконечности.

Асимптоты бывают двух видов: вертикальные и наклонные (в частности, горизонтальные).

Прямая, задаваемая уравнением $x=a$,— *вертикальная* асимптота, если хотя бы один из пределов *)

$$\lim_{x\to a} f(x), \qquad \lim_{x\to a+0} f(x), \qquad \lim_{x\to a-0} f(x)$$

равен $+\infty$ (или $-\infty$).

Примеры. 1. График функции $y=\frac{1}{x}$ имеет вертикальную асимптоту $x=0$, так как

$$\lim_{x\to -0}\frac{1}{x}=+\infty \quad \text{и} \quad \lim_{x\to -0}\frac{1}{x}=-\infty.$$

2. График функции $y=\frac{1}{x^2}$ имеет вертикальную асимптоту $x=0$, так как

$$\lim_{x\to 0}\frac{1}{x^2}=+\infty.$$

3. Вертикальной асимптотой для графика функции $y=\operatorname{tg} x$ будет каждая из вертикальных прямых $x=\pi k$ $(k\in Z)$.

Для существования *наклонной* асимптоты, задаваемой уравнением $y=kx+b$, необходимо и достаточно, чтобы существовали пределы

$$k=\lim\frac{f(x)}{x}, \qquad b=\lim[f(x)-kx]$$

при x, стремящемся к $+\infty$ (или к $-\infty$).

Пример. Найти асимптоты графика функции

$$y=2x-\frac{1}{x}.$$

Вертикальной асимптотой является прямая $x=0$. Коэффициенты уравнения наклонной асимптоты найдем из равенств

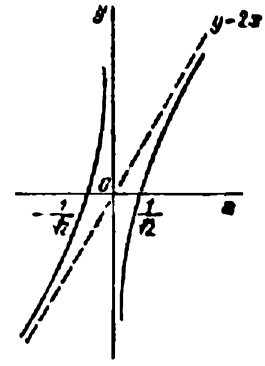


Рис. 9.17

$$k=\lim_{x\to+\infty}\frac{f(x)}{x}=\lim_{x\to+\infty}\frac{2x-\frac{1}{x}}{x}=\lim_{x\to+\infty}\left(2-\frac{1}{x^2}\right)=2,$$

$$b=\lim_{x\to+\infty}[f(x)-kx]=\lim_{x\to+\infty}\left[2x-\frac{1}{x}-2x\right]=\lim_{x\to+\infty}\left(-\frac{1}{x}\right)=0.$$

Таким образом, данная функция имеет наклонную асимптоту, задаваемую уравнением

$$y=2x.$$

График функции $y=2x-\frac{1}{x}$ изображен на рис. 9.17. Пунктирная прямая — наклонная асимптота.

3.6. Построение графика функции. Исследование функции и построение ее графика обычно состоит в следующем:

1. Находят область определения функции.

2. Выясняют, будет ли данная функция периодической.

3. Выясняют, будет ли данная функция четной (или нечетной). Если функция оказывается четной, то ее график симметричен относительно оси Oy прямоугольной декартовой системы координат. Если

*) Второй и третий пределы называют *односторонними* (см. п. 6.2 гл. 8).

функция нечетная, то ее график центрально симметричен относительно начала координат.

4. В случае, если область определения функции представляет собой либо всю числовую прямую, либо содержит лучи $[a; +\infty)$ или $(-\infty; b]$, выясняют поведение функции в бесконечности, вычисляя соответствующие пределы

$$\lim_{x\to+\infty} f(x), \qquad \lim_{x\to-\infty} f(x).$$

5. Находят пересечение графика функции с осью Oy, вычисляя $f(0)$, и с осью Ox, решая уравнение $f(x)=0$.

6. Находят точки максимума и минимума, устанавливая промежутки возрастания и убывания функции.

7. Находят множество значений (область изменения) функции $E(f)$.

8. Находят точки перегиба функции, устанавливая промежутки, где кривая обращена вогнутостью вверх или вниз, причем в точках перегиба вычисляют угол наклона касательной.

9. Находят асимптоты функции.

При исследовании конкретных функций не обязательно придерживаться указанной последовательности исследования функции. Можно даже опускать выяснение тех или иных свойств, если они достаточно очевидны. Так, при исследовании рациональных функций обычно опускают исследование их периодичности по той причине, что из всех элементарных функций лишь тригонометрические функции обладают свойством периодичности.

Пример. Исследовать свойства функции

$$f(x)=\frac{2x^2+3x-4}{x^2}$$

и построить ее график.

1. Область определения функции — вся числовая прямая, за исключением точки $x=0$.

2. Функция не является четной и не является нечетной, так как

$$f(x)\neq f(-x) \quad \text{и} \quad f(x)\neq -f(-x).$$

3. Для того чтобы выяснить поведение функции при неограниченном возрастании (и убывании) аргумента, вычислим пределы

$$\lim_{x\to+\infty} \frac{2x^2+3x-4}{x^2}=2, \qquad \lim_{x\to-\infty} \frac{2x^2+3x-4}{x^2}=2.$$

4. График данной функции не пересекает ось Oy. Точки пересечения графика функции с осью Ox определяются как корни уравнения

$$\frac{2x^2+3x-4}{x^2}=0,$$

которые будут следующими:

$$x_1=\frac{-3+\sqrt{41}}{4}, \qquad x_2=\frac{-3-\sqrt{41}}{4}.$$

5. Найдем точки максимума и минимума функции. Вычислим производную данной функции:

$$f'(x)=\frac{-3x^2+8x}{x^4}=\frac{-3x+8}{x^3}.$$

Производная обращается в нуль при $x=8/3$ и не существует в точке $x=0$, но так как точка $x=0$ не принадлежит области определения данной функции, то она не является критической точкой.

Решим неравенство $f'(x) > 0$:

$$\frac{-3x+8}{x^2} > 0 \Rightarrow x \in \left(0; \frac{8}{3}\right);$$

т. е. функция $f(x)$ возрастает при $x \in \left(0; \frac{8}{3}\right)$ и убывает при $x \in (-\infty; 0) \cup$ $\cup \left(\frac{8}{3}; +\infty\right)$. Следовательно, точка $x = 8/3$ — точка максимума функции.

6. Область изменения функции (множество значений) найдем, используя результаты пп. 3) и 5). Нетрудно проверить, что при всех $x \in (-\infty; 3/4)$ график функции $f(x)$ лежит ниже прямой $y = 2$, так как для этих значений выполняется неравенство

$$f(x) < 2,$$

а при $x \in \left(\frac{3}{4}; +\infty\right)$ график функции лежит выше прямой $y = 2$. Следовательно, максимальное значение функции $f\left(\frac{8}{3}\right) = \frac{41}{16}$ является наибольшим во всей области определения функции.

При x, стремящемся к нулю,

$$\lim_{x \to 0} f(x) = \lim_{x \to 0} \frac{2x^2+3x-4}{x^2} = -\infty.$$

Таким образом, множество значений функции — промежуток

$$\left(-\infty; \frac{41}{16}\right].$$

7) Вычислим вторую производную функции:

$$f''(x) = \frac{6x-24}{x^4}.$$

Вторая производная в промежутке $(4; +\infty)$ положительна, и, следовательно, в этом промежутке график функции направлен вогнутостью вверх. В промежутках $(-\infty, 0)$ и $(0; 4)$ вторая производная отрицательна, и график функции направлен вогнутостью вниз. В точке $x = 4$ график функции имеет перегиб.

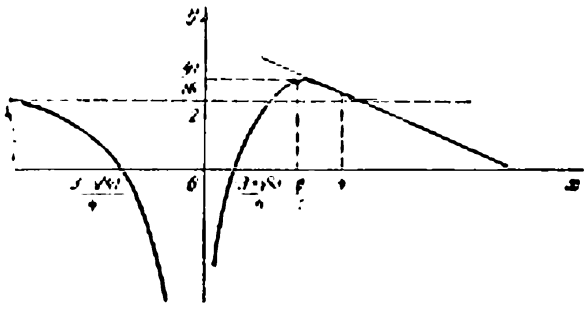


Рис. 9.18

8) Прямая, задаваемая уравнением $x = 0$, — вертикальная асимптота, так как

$$\lim_{x \to 0} \left(\frac{2x^2+3x-4}{x^2}\right) = \lim_{x \to 0} \left(2 + \frac{3}{x} - \frac{4}{x^2}\right) = -\infty.$$

Покажем, что функция имеет и наклонную асимптоту, задаваемую уравнением

$$y = kx + b.$$

Найдем коэффициенты этого уравнения k и b:

$$k=\lim_{x\to\infty}\frac{f(x)}{x}=\lim_{x\to\infty}\frac{2x^2+3x-4}{x^3}=\lim_{x\to\infty}\left(\frac{2}{x}+\frac{3}{x^2}-\frac{4}{x^3}\right)=0,$$

$$b=\lim_{x\to\infty}[f(x)-0\cdot x]=\lim_{x\to\infty}\frac{2x^2+3x-4}{x^2}=\lim_{x\to\infty}\left[2+\frac{3}{x}-\frac{4}{x^2}\right]=2.$$

Значит, данная функция имеет горизонтальную асимптоту $y=2$.

На основании проведенного исследования можно начертить график данной функции (рис. 9.18).

3.7. Экстремумы функции двух независимых переменных. Пусть функция $z=f(x, y)$ определена в открытой области D и точка (x_0, y_0) принадлежит D.

Говорят, что функция $f(x, y)$ имеет в точке $(x_0; y_0)$ *максимум* (*минимум*), если существует такая ε-окрестность точки $(x_0; y_0)$, что для всех точек этой ε-окрестности выполняется неравенство

$$f(x, y)\leqslant f(x_0, y_0)\ (f(x, y)\geqslant f(x_0, y_0)).$$

Если ε-окрестность можно выбрать такой, чтобы знак равенства был исключен, т. е. для любой точки ε-окрестности (кроме точки $(x_0; y_0)$) выполнялось бы строгое неравенство

$$f(x, y)<f(x_0, y_0)\ (f(x, y)>f(x_0, y_0),)$$

то говорят, что в точке $(x_0; y_0)$ функция имеет *собственный максимум* (*минимум*).

Максимум и минимум функции называются *экстремумами*.

Необходимое условие существования экстремума. Если в точке $(x_0; y_0)$ существуют частные производные функции $z=f(x, y)$, то равенство нулю частных производных:

$$f'_x(x_0, y_0)=0,\quad f'_y(x_0, y_0)=0,$$

является необходимым условием существования экстремума.

Точки, в которых частные производные f'_x и f'_y одновременно обращаются в нуль, называют *стационарными*.

Так же как и в случае функции одной переменной, в стационарной точке функция не обязательно имеет экстремум (так же как и в случае функции одной переменной, равенство нулю частных производных первого порядка функции двух переменных является *необходимым*, но не достаточным условием существования экстремума).

Достаточное условие существования экстремума функции. Пусть функция $z=f(x, y)$ определена, непрерывна и имеет частные производные первого и второго порядка в окрестности точки $(x_0; y_0)$, которая является стационарной точкой, т. е.

$$f'_x(x_0, y_0)=0,\quad f'_y(x_0, y_0)=0.$$

Если в стационарной точке $(x_0; y_0)$ выполняется неравенство

$$f''_{xx}\cdot f''_{yy}-(f''_{xy})>0,$$

то в этой точке функция имеет экстремум. Этот экстремум — максимум, если $f''_{xx}<0$, и минимум, если $f''_{xx}>0$. Если же

$$f''_{xx}f''_{yy}-(f''_{xy})^2<0,$$

то в точке $(x_0; y_0)$ функция не имеет экстремума.

Если в стационарной точке $(x_0; y_0)$ выполняется равенство

$$f''_{xx}f''_{yy}-(f''_{xy})^2=0,$$

то для решения вопроса о существовании экстремума необходимо проводить дополнительное исследование функции.

Пример. Исследовать на экстремум функцию

$$z=\sqrt{1-x^2-y^2}.$$

Поверхность, определяемая данным уравнением, представляет собой полусферу (рис. 9.19). Вычисляя частные производные данной функции:

$$\frac{\partial z}{\partial x}=-\frac{x}{\sqrt{1-x^2-y^2}},$$

$$\frac{\partial z}{\partial y}=-\frac{y}{\sqrt{1-x^2-y^2}},$$

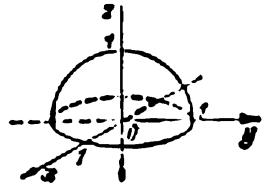


Рис. 9.19

из равенств

$$\frac{\partial z}{\partial x}=0, \quad \frac{\partial z}{\partial y}=0$$

находим, что точка (0; 0) — стационарная точка данной функции.

Вычислим частные производные второго порядка:

$$\frac{\partial^2 z}{\partial x^2}=\frac{y^2-1}{(1-x^2-y^2)^{3/2}}, \quad \frac{\partial^2 z}{\partial x\,\partial y}=-\frac{xy}{(1-x^2-y^2)^{3/2}},$$

$$\frac{\partial^2 z}{\partial y^2}=\frac{x^2-1}{(1-x^2-y^2)^{3/2}}.$$

В точке (0; 0)

$$\frac{\partial^2 z}{\partial x^2}=-1, \quad \frac{\partial^2 z}{\partial x\,\partial y}=0, \quad \frac{\partial^2 z}{\partial y^2}=-1$$

и

$$\frac{\partial^2 z}{\partial x^2}\cdot\frac{\partial^2 z}{\partial y^2}-\left(\frac{\partial^2 z}{\partial x\,\partial y}\right)^2=1>0.$$

Следовательно, в точке (0; 0) функция имеет экстремум; этот экстремум — максимум, так как $\frac{\partial^2 z}{\partial x^2}=-1<0$.

§ 4. Первообразная. Неопределенный интеграл

4.1. Понятия первообразной и неопределенного интеграла. При решении ряда задач механики и физики необходимо восстанавливать функцию по ее известной производной. Так, если материальная точка движется по прямой, не меняя направления своего движения, и расстояние, пройденное точкой,— известная функция времени $s(t)$, то мгновенная скорость точки вычисляется как производная пройденного пути:

$$v=s'(t).$$

Пусть теперь нам известен закон изменения скорости точки, т. е. задана функция $v(t)$, и требуется найти расстояние, которое прошла материальная точка за некоторый промежуток времени T. Для того чтобы решить сформулированную задачу, необходимо по известной производной (в нашем случае скорости $v(t)$) восстановить саму функцию (в нашем случае $s(t)$).

Восстановление функции по известной ее производной составляет одну из основных задач интегрального исчисления. Эта задача является обратной основной задаче дифференциального исчисления, которая состояла в нахождении производной данной функции.

Функция $F(x)$ называется *первообразной* для функции $f(x)$ на данном промежутке числовой оси, если для всех значений x из этого промежутка функция $f(x)$ является производной функции $F(x)$:

$$F'(x)=f(x). \tag{1}$$

Основное свойство первообразных x: *если функция $F(x)$ — первообразная функции $f(x)$, то и функция $F(x)+C$, где C — любая постоянная,— также первообразная функции $f(x)$.* Обратно, *если $F(x)$ — некоторая первообразная функции $f(x)$, то любая первообразная функции $f(x)$ может быть записана в виде*

$$F(x)+C,$$

где C — некоторая постоянная.

Таким образом, для того чтобы найти все первообразные данной функции $f(x)$, достаточно найти лишь одну первообразную $F(x)$; все остальные первообразные будут получаться из найденной первообразной добавлением постоянных слагаемых.

Выражение $F(x)+C$, где C — произвольная постоянная, представляет собой общий вид функции, которая имеет производную $f(x)$. Это выражение называют *неопределенным интегралом* функции $f(x)$ и пишут

$$F(x)+C=\int f(x)\,dx.$$

Произведение $f(x)\,dx$ называется *подынтегральным выражением*, а $f(x)$ — *подынтегральной функцией*.

Нахождение множества всех первообразных функций $f(x)$ называется *интегрированием* этой функции.

Используя формулу (1), из табл. 8 п. 1.5 можно получить следующую таблицу интегралов:

1) $\int 0\cdot dx=C$;

2) $\int 1\cdot dx=x+C$;

3) $\int x^{\alpha}dx=\frac{x^{\alpha+1}}{\alpha+1}+C \qquad (\alpha\neq -1)$;

4) $\int \frac{1}{x}dx=\ln|x|+C$;

5) $\int \frac{1}{1+x^2}dx=\operatorname{arctg}x+C$;

6) $\int \frac{1}{\sqrt{1-x^2}}dx=\arcsin x+C$;

7) $\int a^x dx=\frac{a^x}{\ln a}+C$;

8) $\int \sin x\,dx=-\cos x+C$;

9) $\int \cos x\,dx=\sin x+C$;

10) $\int \frac{1}{\cos^2 x}dx=\operatorname{tg}x+C$;

11) $\int \frac{1}{\sin^2 x}dx=-\operatorname{ctg}x+C$.

4.2. Простейшие правила интегрирования.

1. Пусть $F'(x)$ — производная функции $F(x)$. По определению неопределенного интеграла

$$\int F'(x)\,dx = F(x) + C.$$

2. Постоянный множитель можно выносить из-под знака интеграла:

$$\int c \cdot f(x)\,dx = c \cdot \int f(x)\,dx.$$

3. Интеграл суммы двух функций равен сумме интегралов слагаемых:

$$\int [f(x) \pm g(x)]\,dx = \int f(x)\,dx \pm \int g(x)\,dx.$$

4. Если функция $F(x)$ является первообразной функции $f(x)$, то первообразной функции $f(ax+b)$ является функция $\frac{1}{a}F(ax+b)$ *).

Пример 1. Вычислить неопределенный интеграл

$$\int (6x^3 + x - 1)\,dx.$$

По правилам 2, 3 данный неопределенный интеграл можно представить в виде суммы интегралов:

$$\int (6x^3 + x - 1)\,dx = 6\int x^3\,dx + \int x\,dx - \int 1 \cdot dx.$$

Используя формулу 3) таблицы неопределенных интегралов, получаем

$$\int (6x^3 + x - 1)\,dx = \frac{3}{2}x^4 + \frac{1}{2}x^2 - x + C.$$

Пример 2. Проинтегрировать функцию

$$f(x) = \frac{x - \sqrt{x}}{\sqrt[3]{x}}.$$

Представим данную функцию $f(x)$ в виде

$$\frac{x - \sqrt{x}}{\sqrt[3]{x}} = \frac{x}{\sqrt[3]{x}} - \frac{\sqrt{x}}{\sqrt[3]{x}} = \frac{x}{x^{1/3}} - \frac{x^{1/2}}{x^{1/3}} = x^{2/3} - x^{1/6}.$$

Используя правила 2, 3 и формулу 3) таблицы неопределенных интегралов, получаем

$$\int \frac{x - \sqrt{x}}{\sqrt[3]{x}}\,dx = \int \left[x^{2/3} - x^{1/6}\right] dx = \int x^{2/3}\,dx - \int x^{1/6}\,dx = \frac{3}{5}x^{5/3} - \frac{6}{7}x^{7/6} + C.$$

С помощью правил 2, 3 можно интегрировать лишь очень простые функции. Существуют гораздо более эффективные (но и более сложные) методы интегрирования функций. Такими методами являются метод интегрирования по частям (см. п. 4.3) и метод замены переменной (см. п. 4.4).

4.3. Метод интегрирования по частям. Пусть $u(x)$ и $v(x)$ — две непрерывные функции, имеющие непрерывные производные. Метод

*) Правило 4 есть частный случай правила замены переменной в неопределенном интеграле (см. п. 4.4).

интегрирования по частям основан на следующих двух равенствах:

$$(u\cdot v)'=u'\cdot v+u\cdot v', \quad \int (u\cdot v)'\,dx=u\cdot v+C,$$

первое из которых представляет собой правило дифференцирования произведения двух функций, а второе — правило интегрирования 1 п. 4.2. Из приведенных равенств следует формула интегрирования по частям*)

$$\int u\cdot v'\,dx=u\cdot v-\int u'\cdot v\,dx. \tag{2}$$

Ниже приведены некоторые интегралы, которые могут быть вычислены с помощью формулы (2).

Пример. Вычислить интеграл

$$\int x\cos x\,dx.$$

Вычислим данный интеграл методом интегрирования по частям. Представим подынтегральную функцию в виде произведения двух сомножителей — функции $u(x)=x$ и производной некоторой функции $v(x)$, о которой известно, что

$$v'(x)=\cos x.$$

Производная функции $u(x)$ равна единице:

$$u'(x)=1,$$

а одна из первообразных функции $v'(x)$ равна $\sin x$:

$$v(x)=\int \cos dx=\sin x.$$

Подставляя $u(x)$, $v(x)$ и $u'(x)$ в правую часть равенства (2), получаем

$$\int x\cos x\,dx=x\sin x-\int \sin x\,dx.$$

Таким образом, вычисление данного интеграла сведено к вычислению табличного интеграла

$$\int \sin x\,dx=-\cos x+C$$

и, следовательно,

$$\int x\cos x\,dx=x\sin x+\cos x+C.$$

Очевидно, что в более сложных случаях разложение подынтегральной функции на сомножители может быть не единственным и его следует выбирать таким, чтобы интегрирование функции $v'(x)$ не представляло трудностей и чтобы вычисление интеграла, стоящего в правой части формулы (2), было проще, нежели вычисление интеграла, стоящего в левой части.

Методом интегрирования по частям могут быть, в частности, вычислены интегралы вида

$$\int x^k \ln^m x\,dx, \quad \int x^k \sin bx\,dx, \quad \int x^k e^{ax}\,dx, \quad \int x^k \cos bx\,dx,$$

где k, m — натуральные, а a, b — любые действительные числа.

*) Равенство (2) также может быть записано в виде

$$\int u\,dv=uv-\int v\,du.$$

Пример 1. Вычислить интеграл

$$\int x^3 \ln x \, dx.$$

Положим $u=\ln x$, $v'=x^3$, так что $u'=\frac{1}{x}$, $v=\frac{1}{4}x^4$. Используя формулу интегрирования по частям, получим

$$\int x^3 \ln x \, dx=\frac{1}{4}x^4 \ln x-\int \frac{1}{4}x^4 \cdot \frac{1}{x}\, dx=$$

$$=\frac{1}{4}x^4 \ln x-\frac{1}{4}\int x^3\, dx=\frac{1}{4}x^4 \ln x-\frac{1}{16}x^4+C.$$

Пример 2. Вычислить интеграл

$$\int x^2 \sin x \, dx.$$

Положим $u=x^2$, $v'=\sin x$, так что $u'=2x$, $v=-\cos x$. На основании формулы интегрирования по частям имеем

$$\int x^2 \sin x \, dx=-x^2 \cos x+2\int x \cos x \, dx.$$

Применяя к интегралу, стоящему справа, формулу интегрирования по частям, окончательно получаем

$$\int x^2 \sin x \, dx=-x^2 \cos x+2x \sin x+2 \cos x+C.$$

Пример 3. Вычислить интеграл

$$\int xe^{2x}\, dx.$$

Положим $u=x$, $v'=e^{2x}$. Тогда $u'=1$, $v=\frac{1}{2}e^{2x}$. По формуле (2) исходный интеграл запишется в виде

$$\int xe^{2x}\, dx=\frac{1}{2}xe^{2x}-\frac{1}{2}\int e^{2x}\, dx=\frac{1}{2}xe^{2x}-\frac{1}{4}e^{2x}+C.$$

Так же, последовательным применением интегрирования по частям, вычисляются интегралы

$$\int P(x)\, e^{ax}\, dx, \quad \int P(x) \sin bx \, dx, \quad \int P(x) \cos bx \, dx,$$

где $P(x)$ — многочлен.

Методом интегрирования по частям можно также вычислить интегралы

$$\int e^{ax} \sin bx \, dx, \quad \int e^{ax} \cos bx \, dx.$$

Если к данным интегралам применить метод интегрирования по частям (в обоих случаях положив $v'=e^{ax} \Rightarrow v=\frac{1}{a}e^{ax}$), то получим следующие два равенства:

$$\int e^{ax} \cos bx \, dx=\frac{1}{a}e^{ax} \cos bx+\frac{b}{a}\int e^{ax} \sin bx \, dx, \tag{3}$$

$$\int e^{ax} \sin bx \, dx=\frac{1}{a}e^{ax} \sin bx-\frac{b}{a}\int e^{ax} \cos bx \, dx. \tag{4}$$

Умножая обе части равенства (4) на величину b/a и складывая полученное равенство с равенством (3), получим линейное уравнение относительно интеграла $\int e^{ax}\cos bx\,dx$, в результате решения которого найдем требуемый интеграл:

$$\int e^{ax}\cos bx\,dx=\frac{b\sin bx+a\cos bx}{a^2+b^2}\cdot e^{ax}+C.$$

Аналогично находим и второй интеграл:

$$\int e^{ax}\sin bx\,dx=\frac{a\sin bx-b\cos bx}{a^2+b^2}\cdot e^{ax}+C_1.$$

Пример 4. В качестве еще одного примера применения метода интегрирования по частям получим рекуррентную формулу для вычисления интеграла

$$I_n=\int\frac{dx}{(x^2+a^2)^n}.$$

Полагая

$$u=\frac{1}{(x^2+a^2)^n},\quad v'=1,$$

по формуле (2) имеем

$$I_n=\frac{x}{(x^2+a^2)^n}+2n\int\frac{x^2}{(x^2+a^2)^{n+1}}\,dx. \tag{5}$$

Преобразуем последний интеграл:

$$\int\frac{x^2}{(x^2+a^2)^{n+1}}\,dx=\int\frac{(x^2+a^2)-a^2}{(x^2+a^2)^{n+1}}\,dx=\int\frac{dx}{(x^2+a^2)^n}-a^2\int\frac{dx}{(x^2+a^2)^{n+1}}=$$

$$=I_n-a^2I_{n+1}.$$

Подставляя полученное выражение в равенство (5), получаем

$$I_n=\frac{x}{(x^2+a^2)^n}+2nI_n-2na^2I_{n+1},$$

откуда

$$I_{n+1}=\frac{1}{2na^2}\cdot\frac{x}{(x^2+a^2)^n}+\frac{2n-1}{2n}\cdot\frac{1}{a^2}I_n. \tag{6}$$

Рекуррентная формула (6) позволяет свести вычисление интеграла

$$I_{n+1}=\int\frac{dx}{(x^2+a^2)^{n+1}}$$

к вычислению интеграла I_n.

Вычислим интеграл I_1 с помощью правила 4 (см. п. 4.2) и табличного интеграла 5) (см. п. 4.1):

$$I_1=\int\frac{dx}{x^2+a^2}=\frac{1}{a^2}\int\frac{dx}{\left(\frac{x}{a}\right)^2+1}=\frac{1}{a}\operatorname{arctg}\frac{x}{a}.$$

Зная интеграл I_1, по формуле (6) можно вычислить интеграл I_2. По известному интегралу I_2 с помощью той же самой формулы (6) можно вычислить интеграл I_3 и т. д.

4.4. Интегрирование методом замены переменной. Одним из самых эффективных методов интегрирования является метод замены переменной или метод подстановки. Пусть требуется вычислить интеграл

$$\int f(x)\,dx.$$

Метод замены переменной основан на свойстве инвариантности дифференциала (см. п. 2.3) и заключается в том, что для подынтегрального выражения $f(x)\,dx$ подбирают такую функцию $y=\varphi(x)$, чтобы оно могло быть записано в виде

$$f(x)\,dx = g(\varphi(x))\cdot\varphi'(x)\,dx, \tag{7}$$

где функция $g(y)$ более удобна для интегрирования, нежели функция $f(x)$. На основании свойства инвариантности дифференциала равенство (7) записывается в виде

$$f(x)\,dx = g(y)\,dy, \tag{8}$$

что позволяет свести вычисление данного интеграла к вычислению интеграла

$$\int g(y)\,dy.$$

Вычисляя полученный интеграл

$$\int g(y)\,dy = G(y) + C$$

и возвращаясь к исходной переменной x, получаем

$$\int f(x)\,dx = G(\varphi(x)) + C.$$

Пример 1. Вычислить интеграл

$$\int \sin^2 x \cos x\,dx.$$

Так как $d\sin x = \cos x\,dx$, то, полагая $y=\sin x$, подынтегральное выражение можно преобразовать к виду

$$\sin^2 x \cos x\,dx = \sin^2 x\,d\sin x = y^2\,dy;$$

следовательно,

$$\int \sin^2 x \cos x\,dx = \int y^2\,dy = \frac{1}{3}y^3 + C.$$

Возвращаясь к переменной x, окончательно получаем

$$\int \sin^2 x \cos x\,dx = \frac{1}{3}\sin^3 x + C.$$

Пример 2. Вычислить интеграл

$$\int e^{x^2} x\,dx.$$

Так как $x\,dx = \frac{1}{2}d(x^2)$, то, полагая $y=x^2$, вычисление данного интеграла сводится к вычислению интеграла

$$\frac{1}{2}\int e^y\,dy = \frac{1}{2}e^y + C,$$

следовательно,

$$\int e^{x^2} x\,dx = \frac{1}{2}e^{x^2} + C.$$

Пример 3. Вычислить интеграл

$$\int \operatorname{tg} x\,dx.$$

Подынтегральное выражение запишем в виде

$$\operatorname{tg} x\,dx=\frac{\sin x}{\cos x}\,dx=-\frac{d\cos x}{\cos x}.$$

Обозначая $\cos x=y$, получаем

$$\int \operatorname{tg} x\,dx=\int\frac{dy}{y}=\ln|y|+C=\ln|\operatorname{tg} x|+C.$$

Метод замены переменной упрощает вычисление интегралов вида $\int f(\ln x)\frac{dx}{x}$ (подстановка $y=\ln x$), а также интегралов вида

$$\int f(\sin x)\cos x\,dx,\quad \int f(\cos x)\sin x\,dx$$

(подстановки $y=\sin x$ и $y=\cos x$ соответственно).

4.5. **Некоторые классы интегрируемых функций.** Интегрирование рациональных алгебраических дробей. Использование метода замены переменных и таблицы интегралов (п. 4.1) позволяет провести интегрирование произвольной рациональной функции $\frac{P(x)}{Q(x)}$, где $P(x)$ и $Q(x)$ — многочлены. Интегрирование простейших дробей

$$\frac{A}{x-a}\ \text{и}\ \frac{A}{(x-a)^k},\quad k=2,\ 3,\ \ldots, \tag{9}$$

выполняется непосредственно с помощью табличных интегралов 4) и 3) п. 4.1 соответственно:

$$\int\frac{A}{x-a}\,dx=A\ln|x-a|+C;$$

$$\int\frac{A}{(x-a)^k}\,dx=-\frac{A}{k-1}\cdot\frac{1}{(x-a)^{k-1}}+C.$$

Интегрирование простейших дробей вида

$$\frac{Mx+N}{x^2+px+q}\ \text{и}\ \frac{Mx+N}{(x^2+px+q)^k},\quad k=2,\ 3,\ \ldots, \tag{10}$$

где $\left(\frac{p}{2}\right)^2-q<0$, выполняется с помощью метода подстановки. Полагая $t=x+\frac{p}{2}$, имеем

$$x^2+px+q=t^2+a^2,\quad Mx+N=Mt+\left(N-\frac{Mp}{2}\right),\quad dx=dt,$$

где $a^2=q-\left(\frac{p}{2}\right)^2$. Выполняя указанную подстановку, получаем для первой дроби (10)

$$\int\frac{Mx+N}{x^2+px+q}\,dx=\frac{M}{2}\ln(t^2+a^2)+\frac{1}{a}\left(N-\frac{Mp}{2}\right)\operatorname{arctg}\frac{t}{a}+C, \tag{11}$$

для второй дроби (10) —

$$\int\frac{Mx+N}{(x^2+px+q)^k}=-\frac{1}{k-1}\frac{1}{(t^2+a^2)^{k-1}}+\left(N-\frac{Mp}{2}\right)\int\frac{dt}{(t^2+a^2)^k}, \tag{12}$$

где последний интеграл для любого конкретного значения k можно вычислить по рекуррентной формуле (6) п. 4.3.

Для получения окончательного выражения для интегралов (10) в формулы (11) и (12) следует подставить

$$t=\frac{2x+p}{2} \quad \text{и} \quad a=\sqrt{q-\frac{p^2}{4}}.$$

Так как рациональная алгебраическая дробь может быть записана в виде суммы многочлена и правильной рациональной дроби (см. п. 3.2 гл. 4), а каждая правильная дробь единственным образом разлагается в сумму простейших дробей вида (9), (10) (см. п. 3.3 гл. 4), то интегрирование любой рациональной дроби сводится к интегрированию многочлена и вычислению интегралов дробей (9), (10).

Пример. Вычислить интеграл

$$\int \frac{2x^4-10x^3+7x^2+4x+3}{x^5-2x^3+2x^2-3x+2}\,dx.$$

Подынтегральная функция может быть представлена в виде следующей суммы простейших дробей (см. п. 3.3 гл. 4):

$$\frac{3}{x+2}+\frac{1}{(x-1)^2}-\frac{2}{x-1}+\frac{x-3}{x^2+1},$$

и вычисление данного интеграла сводится к вычислению следующих интегралов:

$$\int \frac{3}{x+2}\,dx=3\ln|x+2|;$$

$$\int \frac{1}{(x-1)^2}\,dx=-\frac{1}{x-1};$$

$$\int \frac{2}{x-1}\,dx=2\ln|x-1|;$$

$$\int \frac{x-3}{x^2+1}\,dx=\int \frac{x\,dx}{x^2+1}-3\int \frac{dx}{x^2+1}.$$

Первый из интегралов, стоящих в правой части последнего равенства, вычислим методом замены переменной: полагая $t=x^2+1$, получаем $dt=2x\,dx$ и

$$\int \frac{x\,dx}{x^2+1}=\frac{1}{2}\int \frac{dt}{t}=\frac{1}{2}\ln|t|=\frac{1}{2}\ln(x^2+1).$$

Так как второй интеграл — табличный интеграл 5), то

$$\int \frac{x-3}{x^2+1}\,dx=\frac{1}{2}\ln(x^2+1)-3\operatorname{arctg}x.$$

Окончательно получаем

$$\int \frac{2x^4-10x^3+7x^2+4x+3}{x^5-2x^3+2x^2-3x+2}\,dx=-\frac{1}{x-1}-3\operatorname{arctg}x+\ln\left|\frac{(x+2)^3\cdot\sqrt{x^2+1}}{(x-1)^2}\right|+C.$$

Подстановки Эйлера. Интегралы вида

$$\int R\left(x,\ \sqrt{ax^2+bx+c}\right)dx,$$

где R — многочлен или рациональная алгебраическая дробь P/Q (P и Q — многочлены) указанных аргументов, сводятся к интегралам от рациональных алгебраических дробей подстановками Эйлера.

Первая подстановка применяется в случае, если $a > 0$. Тогда полагают

$$\sqrt{ax^2+bx+c} = t - \sqrt{a}\,x. \tag{13}$$

Возводя обе части равенства в квадрат, получаем равенство

$$bx + c = t^2 - 2\sqrt{a}\,tx,$$

из которого следует

$$x = \frac{t^2 - c}{2\sqrt{a}\,t + b}. \tag{14}$$

Далее, подставляя в правую часть равенства (13) вместо x его выражение (14), получаем

$$\sqrt{ax^2+bx+c} = \frac{\sqrt{a}\,t^2 + bt + c\sqrt{a}}{2\sqrt{a}\,t + b}. \tag{15}$$

Из равенства (14) также следует

$$dx = 2\cdot\frac{\sqrt{a}\,t^2 + bt + c\sqrt{a}}{(2\sqrt{a}\,t + b)^2}\,dt. \tag{16}$$

Если полученные выражения (14), (15) и (16) подставить в исходный интеграл, то вычисление этого интеграла будет сведено к вычислению интеграла от рациональной алгебраической дроби относительно переменной t.

Пример. Вычислить интеграл

$$\int \frac{dx}{\sqrt{x^2-a^2}}.$$

Полагая $\sqrt{x^2-a^2} = t - x$, получаем

$$x^2 - a^2 = t^2 - 2tx + x^2; \quad 2tx = t^2 + a^2;$$

$$x = \frac{t^2+a^2}{2t}; \quad \sqrt{x^2-a^2} = t - \frac{t^2+a^2}{2t} = \frac{t^2-a^2}{2t};$$

$$dx = \left(\frac{t^2+a^2}{2t}\right)' dt = \frac{2t\cdot 2t - (t^2+a^2)\cdot 2}{4t^2}\,dt = \frac{t^2-a^2}{2t^2}\,dt.$$

Подставим полученные выражения для $\sqrt{x^2-a^2}$ и dx в данный интеграл:

$$\int \frac{dx}{\sqrt{x^2-a^2}} = \int \frac{2t}{t^2-a^2}\cdot\frac{t^2-a^2}{2t^2}\,dt = \int \frac{dt}{t} = \ln|t| + c.$$

Возвращаясь к переменной x, получаем

$$\int \frac{dx}{\sqrt{x^2-a^2}} = \ln\left|x + \sqrt{x^2-a^2}\right| + C.$$

Вторая подстановка применяется, если $c > 0$. В этом случае полагают

$$\sqrt{ax^2+bx+c} = xt + \sqrt{c}.$$

Если возвести обе части этого уравнения в квадрат, то

$$ax^2 + bx + c = x^2t^2 + 2xt\sqrt{c} + c,$$

откуда находим

$$x=\frac{2\sqrt{c}\,t-b}{a-t^2},\quad \sqrt{ax^2+bx+c}=\frac{\sqrt{c}\,t^2-bt+a\sqrt{c}}{a-t^2},$$

$$dx=2\cdot\frac{\sqrt{c}\,t^2-bt+\sqrt{c}\,a}{(a-t^2)^2}\,dt.$$

Подставляя полученные выражения для x, $\sqrt{ax^2+bx+c}$, dx в исходный интеграл, получаем, что вычисление этого интеграла сводится к вычислению интеграла от рациональной алгебраической дроби относительно переменной t.

Пример. Вычислить интеграл

$$\int\frac{dx}{\sqrt{a^2-x^2}},\quad a>0.$$

Полагая $\sqrt{a^2-x^2}=xt+\sqrt{a^2}$, получаем

$$a^2-x^2=x^2t^2+2xt\sqrt{a^2}+a^2;$$

$$x(t^2+1)=-2ta;\quad x=-\frac{2ta}{t^2+1};$$

$$\sqrt{a^2-x^2}=\frac{a(1-t^2)}{t^2+1};\quad dx=\frac{2a(t^2-1)}{(t^2+1)^2}\,dt.$$

Подставим полученные выражения для $\sqrt{a^2-x^2}$ и dx в интеграл:

$$\int\frac{dx}{\sqrt{a^2-x^2}}=\int\frac{(-2)\,dt}{t^2+1}=-2\int\frac{dt}{t^2+1}=-2\operatorname{arctg}t+C.$$

Возвращаясь к переменной x, окончательно получаем

$$\int\frac{dx}{\sqrt{a^2-x^2}}=-2\operatorname{arctg}\frac{\sqrt{a^2-x^2}-a}{x}+C.$$

Третья подстановка пригодна в том случае, когда квадратный трехчлен ax^2+bx+c имеет различные корни x_1 и x_2 и, следовательно, может быть записан в виде

$$ax^2+bx+c=a(x-x_1)(x-x_2).$$

В этом случае полагаем

$$\sqrt{ax^2+bx+c}=t(x-x_1).$$

Возводя обе части уравнения в квадрат, находим

$$x=\frac{-ax_2+x_1t^2}{t^2-a};\quad \sqrt{ax^2+bx+c}=\frac{a(x_1-x_2)\,t}{t^2-a};$$

$$dx=\frac{2a(x_2-x_1)\,t}{(t^2-a)^2}\,dt.$$

Таким образом, и в этом случае вычисление исходного интеграла свелось к вычислению интеграла от рациональной алгебраической дроби.

Интегрирование тригонометрических функций. Интегралы вида

$$\int R(\sin x,\ \cos x)\,dx,$$

где R — многочлен или рациональная алгебраическая дробь указанных переменных, могут быть сведены к интегралам от рациональных

алгебраических дробей подстановкой $t=\operatorname{tg}\frac{x}{2}$, так как

$$\sin x=\frac{2\operatorname{tg}\frac{x}{2}}{1+\operatorname{tg}^2\frac{x}{2}}=\frac{2t}{1+t^2};\quad \cos x=\frac{1-\operatorname{tg}^2\frac{x}{2}}{1+\operatorname{tg}^2\frac{x}{2}}=\frac{1-t^2}{1+t^2};$$

$$x=2\operatorname{arctg} t;\quad dx=\frac{2dt}{1+t^2}.$$

Подстановка $t=\operatorname{tg}\frac{x}{2}$, называемая *универсальной*, довольно часто приводит к сложным вычислениям. Во многих случаях интегралы указанного выше вида могут быть вычислены с помощью более простых подстановок.

Пример. Вычислить интеграл

$$\int \sin^2 x\cos^3 x\,dx.$$

Универсальной подстановкой $t=\operatorname{tg}\frac{x}{2}$ данный интеграл приводится к интегралу

$$\int\frac{8t^2(1-t^2)^3}{(1+t^2)^6}\,dt,$$

вычисление которого требует довольно громоздких вычислений.

Проще данный интеграл может быть вычислен подстановкой

$$t=\sin x;\quad dt=\cos x\,dx;$$

$$\int \sin^2 x\cos^3 x\,dx=\int \sin^2 x(1-\sin^2 x)\cos x\,dx=\int t^2(1-t^2)\,dt=$$

$$=\int (t^2-t^4)\,dt=\frac{1}{3}t^3-\frac{1}{5}t^5+C=\frac{1}{3}\sin^3 x-\frac{1}{5}\sin^5 x+C.$$

Интегрирование обратных тригонометрических функций в ряде случаев удобно проводить методом интегрирования по частям.

Пример. Вычислить интеграл

$$\int \arcsin x\,dx.$$

Положим $u=\arcsin x$, $v'=1$. Тогда

$$u'=\frac{1}{\sqrt{1-x^2}},\quad v=x.$$

На основании формулы (2) получаем

$$\int \arcsin x\,dx=x\arcsin x-\int\frac{x}{\sqrt{1-x^2}}\,dx.$$

Интеграл, стоящий в правой части равенства, может быть вычислен либо с помощью подстановок Эйлера, либо методом замены переменной.

Вычислим его методом замены переменной, полагая

$$t=1-x^2.$$

Тогда $dt=-2x\,dx$ и

$$\int\frac{x}{\sqrt{1-x^2}}\,dx=-\frac{1}{2}\int\frac{dt}{\sqrt{t}}=-\sqrt{t},$$

Возвращаясь к переменной x, получаем

$$\int \frac{x}{\sqrt{1-x^2}}\,dx = -\sqrt{1-x^2}$$

и, следовательно,

$$\int \arcsin x\,dx = x\arcsin x + \sqrt{1-x^2} + C.$$

§ 5. Определенный интеграл

5.1. Задача вычисления площади плоской фигуры. Пусть на некотором промежутке $[a;\ b]$ задана непрерывная функция $f(x)$, принимающая лишь положительные (неотрицательные) значения. Построим график функции $y=f(x)$ в прямоугольной декартовой системе координат Oxy (рис. 9.20).

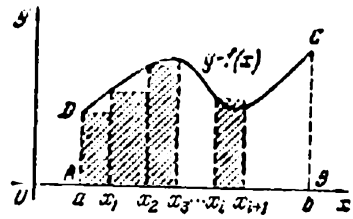


Рис. 9.20

Рассмотрим фигуру $ABCD$, ограниченную графиком функции $y=f(x)$, двумя прямыми, задаваемыми уравнениями $x=a$ и $x=b$, и осью Ox. Эту фигуру принято называть *криволинейной трапецией*.

Разделим основание трапеции AB на n отрезков с концами $x_0=a$, x_1, x_2, ..., x_{n-1}, $x_n=b$ и через точки x_1, x_2, ..., x_{n-1} проведем вертикальные прямые, которые разобьют криволинейную трапецию на n полосок (рис. 9.20).

Для каждой полоски построим прямоугольник, основание которого то же, что и у полоски, а высота совпадает с одной из боковых сторон полоски, скажем с левой (на рис. 9.20 эти прямоугольники заштрихованы).

Основание i-го прямоугольника ($i=0,\ 1,\ 2,\ \ldots,\ n-1$) равно разности $x_{i+1}-x_i$, которую обозначим Δx_i, а высота, согласно принятому выше допущению, равна $y_i=f(x_i)$.

Площадь i-го прямоугольника равна $y_i\cdot\Delta x_i=f(x_i)\cdot\Delta x_i$. Сумма площадей S_n всех прямоугольников равна

$$S_n=f(x_0)\,\Delta x_0+f(x_1)\,\Delta x_1+\ldots+f(x_{n-1})\,\Delta x_{n-1}=\sum_{i=0}^{n-1} f(x_i)\,\Delta x_i.$$

Площадь S_n называют приближенным значением площади криволинейной трапеции $ABCD$.

Точное значение площади S получится как предел

$$S=\lim \sum f(x_i)\,\Delta x_i=\lim \sum y_i\,\Delta x_i \qquad (1)$$

в предположении, что все длины Δx_i одновременно стремятся к нулю при неограниченном увеличении числа точек разбиения отрезка AB.

Для обозначения предельного значения суммы вида (1) Лейбниц ввел символ $\int y\,dx$, где $y\,dx$ — вид слагаемых суммы (1), а $\int$ — стилизованная буква S — начальная буква латинского слова Summa.

5.2. Понятие определенного интеграла. Приведенный в предыдущем пункте способ вычисления площади криволинейной трапеции пригоден лишь для весьма узкого класса функций (например, для

непрерывных функций). Само понятие площади криволинейной трапеции, например, для ограниченных функций, имеющих на промежутке $[a; b]$ конечное число точек разрыва, вводится с помощью понятия определенного интеграла.

Пусть ограниченная функция $f(x)$ задана на некотором промежутке $[a; b]$. Разобьем этот промежуток произвольным образом на n частей, вставив между точками a и b точки деления $x_1, x_2, \ldots, x_{n-1}$, причем

$$a = x_0 < x_1 < x_2 < \ldots < x_i < x_{i+1} < \ldots < x_n = b.$$

Наибольшую из разностей $\Delta x_i = x_{i+1} - x_i$ $(i = 0, 1, 2, \ldots, n-1)$ обозначим через λ.

Возьмем в каждом из промежутков $[x_i; x_{i+1}]$ произвольную точку ξ_i: $x_i \leqslant \xi_i \leqslant x_{i+1}$ и составим сумму, называемую *интегральной суммой*:

$$\sigma_n = \sum_{i=0}^{n-1} f(\xi_i)\,\Delta x_i, \tag{2}$$

соответствующую данному разбиению отрезка $[a; b]$. Сумма (2) зависит от выбора точек разбиения x_i отрезка $[a; b]$ и от выбора точек ξ_i.

Говорят, что сумма σ_n имеет предел I при λ, стремящемся к нулю, если для любого сколь угодно малого числа $\varepsilon > 0$ существует такое число $\delta > 0$, что неравенство

$$|\sigma_n - I| < \varepsilon$$

выполняется при любом $\lambda < \delta$ независимо от выбора чисел ξ_i (т. е. для любого разбиения промежутка $[a; b]$ на части, меньшие чем δ, и при любом выборе точек ξ_i).

Конечный предел I суммы σ_n при λ, стремящемся к нулю, называется *определенным интегралом* функции $f(x)$ на промежутке от a до b и обозначается

$$\int_a^b f(x)\,dx. \tag{3}$$

Если такой предел I существует, то говорят, что функция $f(x)$ *интегрируема* на промежутке $[a; b]$.

В записи (3) функция $f(x)$ называется *подынтегральной функцией*, $f(x)\,dx$ — *подынтегральным выражением*, а числа a и b называются соответственно *нижним* и *верхним пределами интеграла*.

Если функция $f(x)$ положительна на промежутке $[a; b]$, то существование конечного предела I (определенного интеграла (3)) означает, что площадь криволинейной трапеции, ограниченной графиком функции $y = f(x)$, осью Ox и прямыми $x = a$ и $x = b$, существует и равна I.

Примером неинтегрируемой функции является функция Дирихле $D(x)$, задаваемая на всей числовой оси условиями:

$$D(x) = \begin{cases} 1, & \text{если } x \text{ рационально,} \\ 0, & \text{если } x \text{ иррационально.} \end{cases}$$

Покажем, что функция $D(x)$ не интегрируема, например, на промежутке $[0; 1]$. Разобьем этот промежуток произвольным образом на n частей. На каждом из промежутков $[x_i; x_{i+1}]$ выберем рациональную точку ξ_i' и иррациональную точку ξ_i'' и составим две интегральные суммы σ_n' и σ_n'', соответствующие

данному разбиению:

$$\sigma_n = \sum_{i=0}^{n-1} D(\xi_i)\,\Delta x_i, \qquad \sigma_n' = \sum_{i=0}^{n-1} D(\xi_i')\,\Delta x_i.$$

Для точек ξ_i и ξ_i' по определению функции $D(x)$ имеем $D(\xi_i)=1$ и $D(\xi_i')=0$ и, следовательно, $\sigma_n=1$, $\sigma_n'=0$.

Аналогичные интегральные суммы σ и σ' можно составлять для любого другого разбиения отрезка $[0; 1]$, выбирая каждый раз ξ_i рациональным, ξ_i' — иррациональным и каждый раз получая $\sigma=1$, $\sigma'=0$.

Итак, можно построить две последовательности интегральных сумм — одну, состоящую только из единиц, и другую, состоящую только из нулей. Пределом первой последовательности будет единица, а пределом второй — нуль. Но так как для существования предела интегральных сумм необходимо, чтобы последовательность интегральных сумм сходилась к числу 1 **независимо** от выбора точек ξ_i, то установленное противоречие и доказывает, что функция Дирихле не интегрируема на промежутке $[0; 1]$.

Свойства определенных интегралов.

1. Если функция $f(x)$ интегрируема на промежутке $[a; b]$, то и функция $k\cdot f(x)$ (k постоянная) также интегрируема на этом промежутке, причем

$$\int_a^b k\cdot f(x)\,dx = k\int_a^b f(x)\,dx,$$

т. е. постоянный множитель можно выносить за знак интеграла.

2. Если функции $f(x)$ и $g(x)$ интегрируемы на промежутке $[a; b]$, то и функции $f(x)\pm g(x)$ также интегрируемы на этом промежутке; при этом

$$\int_a^b [f(x)\pm g(x)]\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx.$$

3. При перестановке пределов интегрирования интеграл меняет знак на противоположный:

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

Из этого равенства, в частности, следует, что

$$\int_a^a f(x)\,dx = 0.$$

4. Если функция $f(x)$ интегрируема на промежутке $[a; b]$ и точка $c\in(a; b)$, то функция интегрируема на промежутках $[a; c]$ и $[c; b]$, причем

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

5. Если неотрицательная функция $f(x)$ интегрируема на промежутке $[a; b]$, то

$$\int_a^b f(x)\,dx \geqslant 0.$$

6. Если функции $f(x)$ и $g(x)$ интегрируемы на промежутке $[a; b]$ и при всех $x \in [a; b]$

$$f(x) \leqslant g(x),$$

то

$$\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx.$$

7. Если функция интегрируема на промежутке $[a; b]$, где $a < b$, и при всех $x \in [a; b]$ справедливо неравенство

$$m \leqslant f(x) \leqslant M,$$

то

$$m(b-a) \leqslant \int_a^b f(x)\,dx \leqslant M(b-a).$$

5.3. Определенный интеграл как функция верхнего предела. Основная формула интегрального исчисления. Если функция $f(t)$ интегрируема на промежутке $[a; b]$, то по свойству 4 определенного интеграла (см. п. 5.2) она интегрируема и в промежутке $[a; x]$, где x — любое число из промежутка $[a; b]$. Заменив верхний предел b определенного интеграла переменной x, получим выражение

$$\Phi(x) = \int_a^x f(t)\,dt, \tag{4}$$

которое можно рассматривать как функцию переменной x. Эта функция обладает следующими свойствами:

1) если функция f интегрируема в промежутке $[a; b]$, то функция $\Phi(x)$ — непрерывная функция в промежутке $(a; b)$;

2) если функция f непрерывна в промежутке $[a; b]$, то в любой внутренней точке этого промежутка функция $\Phi(x)$ имеет производную, равную $f(x)$:

$$\Phi'(x) = f(x).$$

Последнее свойство функции $\Phi(x)$ является одним из важнейших результатов математического анализа.

Для функции $f(x)$, непрерывной в промежутке $[a; b]$, всегда существует первообразная; этой первообразной является определенный интеграл (4) с переменным верхним пределом.

Из этой теоремы, в частности, следует формула

$$\int_a^b f(x)\,dx = F(b) - F(a), \tag{5}$$

где $F(b)$ и $F(a)$ — значения произвольной первообразной функции $f(x)$ в точках $x = b$ и $x = a$ соответственно.

Формулу (5) называют основной формулой интегрального исчисления: *значение определенного интеграла равно разности двух значений любой первообразной функции, вычисленных при* $x=b$ *и при* $x=a$.

Разность, стоящую в правой части формулы (5), обычно обозначают

$$F(x)\Big|_a^b,$$

а саму формулу (5) записывают в виде

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b.$$

Основная формула интегрального исчисления позволяет сводить вычисление определенного интеграла к вычислению значений первообразной (неопределенного интеграла) с использованием всего развитого аппарата вычислений неопределенных интегралов.

Пример. Вычислить определенный интеграл

$$\int_0^{\pi/2} \sin x\,dx.$$

Первообразной функции $\sin x$ будет функция $-\cos x$. По основной формуле интегрального исчисления

$$\int_0^{\pi/2} \sin x\,dx = -\cos x\Big|_0^{\pi/2} = -\cos\frac{\pi}{2} - (-\cos 0) = -0+1=1.$$

Аналогичным образом по основной формуле интегрального исчисления вычисляются следующие интегралы:

$$\int_0^1 x^2\,dx = \frac{1}{3}x^3\Big|_0^1 = \frac{1}{3}-0=\frac{1}{3},$$

$$\int_{1/2}^{5} \frac{1}{x}\,dx = \ln x\Big|_{1/2}^{5} = \ln 5 - \ln\frac{1}{2} = \ln 10,$$

$$\int_{-1}^{1} e^x\,dx = e^x\Big|_{-1}^{1} = e^1 - e^{-1} = e-\frac{1}{e}.$$

5.4. Вычисление площадей и объемов с помощью определенного интеграла. Вычисление площади плоской фигуры. Площадь криволинейной трапеции, ограниченной графиком непрерывной и неотрицательной функции $y=f(x)$, осью Ox и прямыми $x=a$ и $x=b$ (рис. 9.21), вычисляется по формуле

$$S = \int_a^b f(x)\,dx.$$

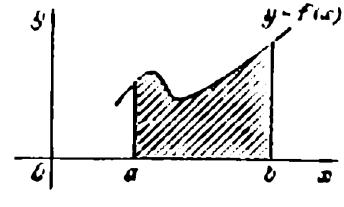


Рис. 9.21

Пример 1. Вычислить площадь фигуры, ограниченной графиком функции $y=e^{2x}$, осью Ox и прямыми $x=0$ и $x=1$ (рис. 9.22). Площадь данной фигуры равна

$$\int_0^1 e^{2x}\,dx=\frac{1}{2}e^{2x}\Big|_0^1=\frac{1}{2}(e^2-e^0)=\frac{1}{2}(e^2-1).$$

Пример 2. Вычислить площадь криволинейной трапеции, ограниченной графиком функции $y=\ln x$, осью Ox и прямыми $x=2$ и $x=4$ (рис. 9.23)

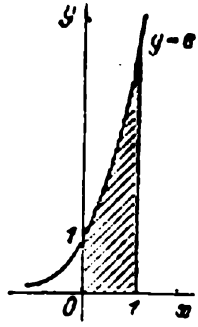


Рис. 9.22

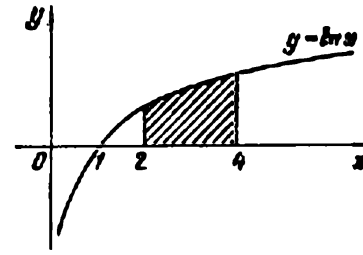


Рис. 9.23

Площадь данной фигуры равна значению интеграла

$$\int_2^4 \ln x\,dx.$$

Найдем первообразную $F(x)$ функции $\ln x$ методом интегрирования по частям, обозначая $\ln x=u$, $v'=1$. Тогда $u'=\frac{1}{x}$ и $v=x$. По формуле интегрирования по частям (см. п. 4.3) получаем

$$F(x)=\int \ln x\,dx=x\ln x-\int x\cdot\frac{1}{x}\,dx=x\ln x-x=$$

$$=x(\ln x-1)=x\ln\frac{x}{e}.$$

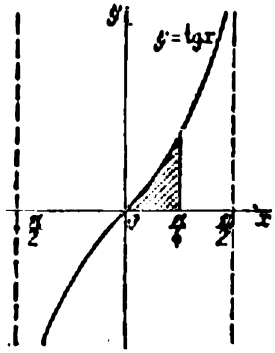


Рис. 9.24

По основной формуле интегрального исчисления (см. п. 5.3)

$$\int_2^4 \ln x\,dx=F(x)\Big|_2^4=x\ln\frac{x}{e}\Big|_2^4=4\ln\frac{4}{e}-2\ln\frac{2}{e}=$$

$$=\ln\frac{256}{e^4}-\ln\frac{4}{e^2}=\ln\frac{64}{e^2}.$$

Пример 3. Вычислить площадь фигуры, ограниченной графиком функции $y=\operatorname{tg}x$, осью Ox и прямой $x=\pi/4$. Площадь искомой фигуры заштрихована на рис. 9.24 и равна значению интеграла

$$\int_0^{\pi/4}\operatorname{tg}x\,dx.$$

Вычислим первообразную $F(x)$ функции $\operatorname{tg} x$ методом замены переменной (см. п. 4.4), обозначая $\cos x = t$. Тогда $dt = -\sin x\,dx$ и

$$F(x) = \int \operatorname{tg} x\,dx = \int \frac{\sin x\,dx}{\cos x} = -\int \frac{dt}{t} = -\ln|t| = -\ln(\cos x).$$

(В последнем выражении знак абсолютной величины опущен, так как $\cos x$ при $0 \leqslant x \leqslant \pi/4$ положителен.)

Далее по основной формуле интегрального исчисления получаем значение искомой площади

$$\int_0^{\pi/4} \operatorname{tg} x\,dx = -\ln(\cos x)\Big|_0^{\pi/4} = \ln(\cos 0) - \ln\left(\cos\frac{\pi}{4}\right) = 0 - \ln\frac{\sqrt{2}}{2} = \ln\sqrt{2}.$$

Вычисление площади поверхности вращения и объема. Площадь поверхности, полученной в результате вращения вокруг оси Ox криволинейной трапеции, ограниченной графиком

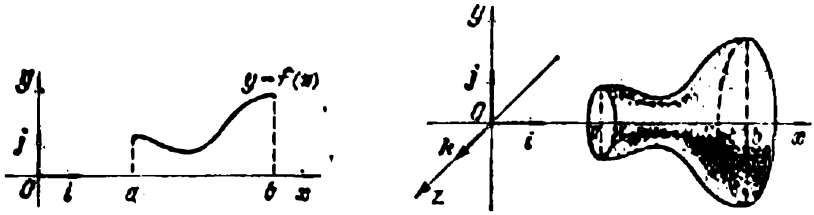


Рис. 9.25

непрерывной неотрицательной функции $y = f(x)$ и прямыми $x = a$ и $x = b$ (рис. 9.25), вычисляется по формуле

$$S = 2\pi \int_a^b f(x)\sqrt{1 + (f'(x))^2}\,dx, \tag{7}$$

а объем — по формуле

$$V = \pi \int_a^b [f(x)]^2\,dx. \tag{8}$$

Пример. Вычислить объем и площадь поверхности шара радиуса R. Шар может быть получен в результате вращения полукруга вокруг оси, содержащей диаметр. Площадь его поверхности и объем можно вычислить по формулам (7) и (8). Уравнение полуокружности с центром в точке $(R; 0)$ и диаметром, принадлежащим оси Ox (рис. 9.26), имеет вид

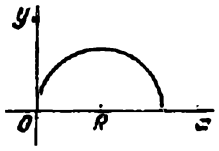


Рис. 9.26

$$y = \sqrt{R^2 - (x-R)^2}. \tag{9}$$

Объем шара, полученного в результате вращения полукруга, ограниченного окружностью (9), равен

$$V = \pi \int_0^{2R} y^2\,dx = \pi \int_0^{2R} [R^2 - (x-R)^2]\,dx = \pi \int_0^{2R} (2Rx - x^2)\,dx =$$

$$= \pi \int_0^{2R} 2Rx\,dx - \pi \int_0^{2R} x^2\,dx = \pi R x^2\Big|_0^{2R} - \frac{\pi}{3}x^3\Big|_0^{2R} = 4\pi R^3 - \frac{8\pi R^3}{3} = \frac{4}{3}\pi R^3.$$

Площадь поверхности шара (площадь сферы) S вычислим по формуле (7):

$$f'(x)=y'(x)=-\frac{(x-R)}{\sqrt{R^2-(x-R)^2}}; \qquad \sqrt{1+(f'(x))^2}=\frac{R}{\sqrt{R^2-(x-R)^2}};$$

$$S=2\pi\int_0^{2R}\sqrt{R^2-(x-R)^2}\,\frac{R}{\sqrt{R^2-(x-R)^2}}\,dx=2\pi\int_0^{2R}R\,dx=2\pi Rx\Big|_0^{2R}=4\pi R^2.$$

§ 6. Дифференциальные уравнения

6.1. Понятие дифференциального уравнения. *Обыкновенным дифференциальным уравнением* называется уравнение, в котором неизвестной является функция одной независимой переменной, причем уравнение содержит не только неизвестную функцию, но и ее производные различных порядков.

Порядок наивысшей производной, входящей в уравнение, называют *порядком* дифференциального уравнения.

Обыкновенные дифференциальные уравнения имеют большое прикладное значение — они широко используются в механике, астрономии, физике и других науках и с их помощью количественно достаточно точно описываются многие явления и процессы, происходящие в природе. Так, например, открытое в начале нашего века физическое свойство радиоактивного вещества — свойство распада — обычно формулируют в виде «закона радиоактивного распада»: скорость радиоактивного распада вещества пропорциональна количеству этого вещества.

Запишем закон радиоактивного распада в виде уравнения для массы радиоактивного вещества. Так как, в согласии с законом радиоактивного распада, масса вещества меняется со временем, то обозначим ее $m(t)$ (t — время). Пусть t_0 — данный момент времени. За время Δt масса радиоактивного вещества уменьшится на величину

$$\Delta m = m(t_0+\Delta t)-m(t_0).$$

Отношение

$$\frac{\Delta m}{\Delta t}=\frac{m(t_0+\Delta t)-m(t_0)}{\Delta t}$$

показывает, насколько изменилась масса радиоактивного вещества за единицу времени, и представляет собой «среднюю скорость» изменения массы.

Но так как при радиоактивном распаде масса вещества (функция $m(t)$) меняется со временем, то отношение $\Delta m/\Delta t$ является функцией Δt, и для точной записи сформулированного выше закона радиоактивного распада необходимо взять предельное значение отношения $\Delta m/\Delta t$ при Δt, стремящемся к нулю:

$$\lim_{\Delta t\to 0}\frac{\Delta m}{\Delta t}.$$

По определению этот предел есть производная функция $m(t)$ по переменной t:

$$\lim_{\Delta t\to 0}\frac{\Delta m}{\Delta t}=m'(t_0).$$

Так как закон радиоактивного распада справедлив в любой момент времени (а не только в момент времени t_0), то математически он может быть записан в виде уравнения

$$m'(t) = -\beta m(t).$$

Здесь β — постоянный положительный коэффициент пропорциональности, зависящий от вида радиоактивного вещества. Знак минус в полученном уравнении указывает на то, что со временем масса радиоактивного вещества убывает, превращаясь в излучение ($m' < 0$ при любом t).

Дифференциальным уравнением первого порядка может быть также описан процесс нагревания (охлаждения) тела.

Пусть известно, что в момент времени t_0 тело имело температуру T_0. Поместим тело в какую-либо среду, имеющую температуру $\theta \neq T_0$. Если $\theta > T_0$, то тело будет нагреваться, если же $\theta < T_0$, то — охлаждаться. В некоторых случаях можно положить, что скорость изменения температуры тела T пропорциональна разности температур, т. е. удовлетворяет дифференциальному уравнению

$$T'(k) = -k(T - \theta), \tag{1}$$

где k — некоторая положительная постоянная. Знак минус указывает на то, что при $T_0 > \theta$ тело будет охлаждаться, а при $T_0 < \theta$ — нагреваться.

Необходимо отметить, что моделирование указанного физического процесса уравнением (1) является довольно грубым и основано на следующих предположениях: считается, что в любой момент времени тело нагрето равномерно по объему, хотя теплообмен происходит через поверхность тела; считается, что температура окружающей среды постоянна и равна начальной температуре среды, несмотря на то что физически очевидно, что температура среды также будет меняться, и т. д. Более точно процесс нагревания тела описывается с помощью дифференциальных уравнений в частных производных.

6.2. Дифференциальные уравнения первого порядка. *Дифференциальным уравнением первого порядка* называют уравнение, в которое входит лишь первая производная неизвестной функции. Это уравнение может быть записано в виде

$$F(x,\ y(x),\ y'(x)) = 0, \tag{2}$$

где x — независимая переменная, $y(x)$ — неизвестная функция, а F — заданная функция трех переменных.

Уравнение (2) связывает три переменные величины x, y и y'. В тех случаях, когда уравнение (2) определяет переменную y' как функцию переменных x, y:

$$y' = f(x,\ y), \tag{3}$$

где f — известная функция, определенная в некоторой области D координатной плоскости Oxy, дифференциальное уравнение называется *разрешенным относительно производной*.

Решением дифференциального уравнения (3) называется такая функция независимой переменной x

$$y = \varphi(x),$$

определенная на некотором промежутке $(a;\ b)$, которая при подстановке в уравнение (3) обращает уравнение в тождество на всем промежутке $(a;\ b)$. Подстановка $y = \varphi(x)$ в уравнение (3) возможна лишь

тогда, когда функция $\varphi(x)$ дифференцируема на всем промежутке $(a; b)$. Кроме того, также необходимо, чтобы при любом значении $x_0 \in (a; b)$ точка с координатами $(x_0; \varphi(x_0))$ принадлежала области определения функции $f(x, y)$.

Решение дифференциального уравнения $y = \varphi(x)$ геометрически можно изобразить на координатной плоскости Oxy в виде линии, задаваемой уравнением $y = \varphi(x)$. Эта линия называется *интегральной кривой* дифференциального уравнения (3).

В теории дифференциальных уравнений центральное место занимают следующие вопросы.

1. Каким условиям должна удовлетворять функция $f(x, y)$, стоящая в левой части уравнения (3), чтобы это дифференциальное уравнение имело решение?

2. Какие дополнительные условия необходимо наложить на функцию $f(x, y)$, чтобы выделить одно-единственное решение из бесконечного множества решений уравнения (3)?

Ответ на эти вопросы дает

Теорема о существовании и единственности решения: *Пусть*

$$y' = f(x, y) \tag{4}$$

— дифференциальное уравнение первого порядка, разрешенное относительно производной, и функция f определена в некоторой открытой области D координатной плоскости Oxy. Если функция $f(x, y)$ непрерывна в области D и ее частная производная $\partial f/\partial y$ — также непрерывная функция в области D, то:

1) *для любой точки $(x_0; y_0) \in D$ найдется решение $y = \varphi(x)$ уравнения* (4), *удовлетворяющее условию*

$$\varphi(x_0) = y_0; \tag{5}$$

2) *если $y = \varphi(x)$ и $y = \psi(x)$ — два решения уравнения* (4), *удовлетворяющие условию*

$$\varphi(x_0) = \psi(x_0),$$

то эти решения тождественно равны при всех тех значениях переменной x, при которых они оба определены *).

Числа x_0, y_0 называются *начальными значениями для решения* $y = \varphi(x)$, а равенство (5) — *начальным условием*. Также говорят, что решение $y = \varphi(x)$ удовлетворяет начальному условию (5) или что решение $y = \varphi(x)$ имеет начальные значения x_0, y_0.

Пример. Решить уравнение, описывающее процесс радиоактивного распада:

$$m' = -\beta \cdot m. \tag{6}$$

Переходя к обозначениям, используемым в настоящем пункте, запишем это уравнение в виде

$$y' = -\beta y. \tag{6'}$$

Правая часть уравнения $f(x, y) = -\beta y$ зависит лишь от переменной y и определена при любом значении y (т. е. на всей координатной плоскости Oxy). Как сама функция $f(x, y) = -\beta y$, так и ее частная производная $\dfrac{\partial f(t, x)}{\partial y} = -\beta$

*) Для существования решения уравнения (4) достаточно только непрерывности функции $f(x, y)$ в области D, но при этом решение может быть не единственным.

Теорема существования и единственности решения допускает и другие формулировки: вместо условия непрерывности частной производной $\partial f/\partial y$ выставляют некоторые другие ограничения на функцию $f(x, y)$.

являются непрерывными функциями переменных x и y на всей координатной плоскости xOy, и, следовательно, правая часть уравнения (6′) удовлетворяет условию теоремы существования и единственности решения.

Непосредственной подстановкой в уравнение (6) можно проверить, что каждая функция *)

$$y=Ce^{-\beta x}, \tag{7}$$

где C — произвольное действительное число, является решением уравнения (6′).

Придавая C всевозможные числовые значения, получим все решения уравнения (6′). Действительно, пусть $y=\varphi(x)$ — произвольное решение уравнения (6′). Покажем, что при подходящем выборе значения C при всех значениях $x \in \mathbb{R}$ справедливо равенство

$$\varphi(x)=Ce^{-\beta x},$$

т. е. функция $\varphi(x)$ тождественно совпадает с функцией $Ce^{-\beta x}$.

Пусть x_0 — некоторая точка и $y_0=\varphi(x_0)$. Положим $C=y_0e^{\beta x_0}$. Тогда решения $y=\varphi(x)$ и $y=Ce^{-\beta x}=y_0e^{-\beta(x-x_0)}$ при $x=x_0$ имеют одинаковые начальные условия (x_0, y_0). Но так как правая часть уравнения (6′) удовлетворяет условию теоремы существования и единственности решения, то эти решения должны совпадать, т. е. $\varphi(x)=y_0e^{-\beta(x-x_0)}$ при всех значениях x. Следовательно, формула (7) задает все решения уравнения (6′).

Возвращаясь к старым переменным (m и t), получаем, что решение уравнения радиоактивного распада вещества определяется формулой

$$m(t)=Ce^{-\beta t}.$$

Для определения константы C достаточно указать какие-либо начальные условия. Например, если известно, что в момент времени $t=t_0$ масса вещества была равна m_0, то константа C будет равна

$$C=m_0e^{\beta t_0},$$

и решение уравнения (6) с начальными условиями (t_0, m_0) имеет вид

$$m(t)=m_0e^{-\beta(t-t_0)}.$$

В уравнении (6) скорость распада характеризуется величиной β. Часто вместо величины β скорость распада характеризуют так называемым периодом полураспада, т. е. временем, за которое масса радиоактивного вещества уменьшается в два раза. Обозначая период полураспада через T_0, получим, что коэффициент β и период полураспада T_0 связаны равенством

$$T_0=\frac{1}{\beta}\ln 2.$$

6.3. Некоторые дифференциальные уравнения первого порядка. Уравнения с разделяющимися переменными. Уравнения вида

$$y'=f(x)\,g(y)$$

называются уравнениями с *разделяющимися* переменными.

Для того чтобы найти решения уравнения с разделяющимися переменными, необходимо обе части уравнения умножить (или разделить) на такое выражение, чтобы в одну часть уравнения входила только переменная x, а в другую — только переменная y, а затем проинтегрировать обе части уравнения. Записывая данное уравнение

*) О методе нахождения этой функции см. в п. 6.3 «Уравнения с разделяющимися переменными», пример 1.

в виде

$$\frac{dy}{dx}=f(x)\,g(y),$$

получаем

$$\frac{dy}{g(y)}=f(x)\,dx \quad \text{и} \quad \int\frac{dy}{g(y)}=\int f(x)\,dx.$$

Решая уравнение с разделяющимися переменными, следует иметь в виду, что при делении обеих частей уравнения на выражение, содержащее x и y, могут быть «потеряны» решения, обращающие это выражение в нуль.

Пример 1. Найти решение уравнения радиоактивного распада (см. уравнение (6′))

$$y'=-\beta y.$$

Запишем уравнение в виде

$$\frac{dy}{dx}=-\beta y.$$

Разделив обе части уравнения на y и умножив на dx, получаем уравнение с разделяющимися переменными, которое можно проинтегрировать:

$$\frac{dy}{y}=-\beta\,dx, \qquad \int\frac{dy}{y}=-\beta\int dx, \qquad \ln|y|=-\beta x+C_0.$$

Полагая постоянную интегрирования C_0 равной $\ln C$ и учитывая, что по физическому смыслу задачи $y>0$, находим

$$y=Ce^{-\beta x}.$$

2. Решить уравнение

$$xyy'+1=y.$$

Запишем уравнение в виде

$$xy\frac{dy}{dx}=y-1.$$

Последовательно разделив обе части уравнения на $x(y-1)$ и умножив на dx, получим уравнение

$$\frac{y\,dy}{y-1}=\frac{dx}{x}.$$

Переменные разделены. Проинтегрируем обе части уравнения:

$$\int\frac{y\,dy}{y-1}=\int\frac{dx}{x}; \qquad \int dy+\int\frac{dy}{y-1}=\int\frac{dx}{x};$$

$$y+\ln|y-1|=\ln|x|+C.$$

Последнее равенство, устанавливающее вид зависимости между переменными x, y, и представляет собой решение данного дифференциального уравнения.

При делении мы предполагали, что $x\neq 0$ и $y\neq 1$, т. е. могли «потерять» решения $x=0$ и $y=1$. Подставляя $x=0$ и $y=1$ в исходное уравнение, убеждаемся, что $y=1$ является его решением, а $x=0$ — нет.

Уравнения вида

$$y'=f(ax+by), \quad a \text{ и } b \text{ — числа},$$

приводятся к уравнению с разделяющимися переменными заменой

$$z=ax+by,$$

где переменная z считается функцией x.

Пример. Решить уравнение

$$y' - y = 2x - 3.$$

Записывая уравнение в виде

$$y' = y + 2x + 3$$

и производя замену

$$z = y + 2x \Rightarrow z' = y' + 2 \Rightarrow y' = z' - 2,$$

получаем уравнение с разделяющимися переменными

$$z' = z + 5, \quad \text{или} \quad \frac{dz}{dx} = z + 5, \quad \text{или} \quad \frac{dz}{z+5} = dx.$$

Проинтегрировав, получаем

$$\ln|z+5| = x + C,$$
$$\ln|y + 2x + 5| = x + C.$$

Нетрудно проверить, что $z = -5$ (или $y = -2x - 5$) также является решением данного уравнения.

Однородные уравнения. Дифференциальные уравнения, которые могут быть записаны в виде

$$y' = f\left(\frac{y}{x}\right),$$

называются *однородными*. Однородное уравнение сводится к уравнению с разделяющимися переменными заменой

$$y = tx,$$

где новая переменная t рассматривается как функция x.

Пример. Решить уравнение

$$2x^3y' = 2x^2y - y^3.$$

Разделим обе части уравнения на $2x^3$:

$$y' = \frac{y}{x} - \frac{1}{2}\frac{y^3}{x^3}.$$

Полученное уравнение однородное. Сделаем замену

$$y = tx \Rightarrow y' = t'x + t.$$

Теперь однородное уравнение записывается в виде

$$t'x + t = t - \frac{1}{2}t^3; \quad t'x = -\frac{1}{2}t^3;$$

$$\frac{dt}{dx}x = -\frac{1}{2}t^3; \quad -\frac{2dt}{t^3} = \frac{dx}{x}.$$

Проинтегрировав последнее уравнение, находим решение данного дифференциального уравнения:

$$t^{-2} = \ln|x| + C \quad \text{или} \quad \left(\frac{x}{y}\right)^2 = \ln|x| + C.$$

Линейные уравнения. Дифференциальное уравнение

$$y' + a(x)\,y = b(x),$$

где $a(x)$ и $b(x)$ — некоторые функции, называется *линейным*.

Решение линейного уравнения ищется в два этапа.

1. Сначала решают уравнение с разделяющимися переменными

$$y' + a(x)\,y = 0.$$

2. В полученном решении этого уравнения произвольную постоянную интегрирования C заменяют на неизвестную функцию $C(x)$. Выражение, полученное для y после такой замены, подставляют в исходное уравнение и находят функцию $C(x)$.

Пример. Решить уравнение

$$y' + \frac{1}{x} y = \frac{1}{x^2}.$$

1. Решаем уравнение

$$y' + \frac{1}{x} y = 0; \quad \frac{dy}{dx} = -\frac{y}{x}; \quad \frac{dy}{y} = -\frac{dx}{x};$$

$$\int \frac{dy}{y} = -\int \frac{dx}{x}. \quad \ln|y| = \ln\left|\frac{1}{x}\right| + \ln|C|.$$

(Здесь постоянную интегрирования удобно записать как $\ln|C|$.) Последнее равенство можно преобразовать к виду

$$\ln|yx| = \ln|C|; \quad |yx| = |C|. \quad |y| = \left|\frac{C}{x}\right|.$$

Далее знак абсолютной величины можно опустить, так как C — произвольная постоянная, которая может быть как положительной, так и отрицательной.

2. Заменяя произвольную постоянную C на $C(x)$, имеем

$$y = \frac{C(x)}{x}; \quad y' = \frac{C'(x) \cdot x - C(x)}{x^2}.$$

Подставляя y и y' в исходное уравнение, получаем уравнение для нахождения $C(x)$:

$$\frac{C'}{x} - \frac{C}{x^2} + \frac{C}{x^2} = \frac{1}{x^2};$$

$$C' = \frac{1}{x}; \quad C = \ln|x| + C_1,$$

где C_1 — произвольная постоянная.

Итак, окончательно решение данного уравнения имеет вид

$$y = \frac{\ln|x|}{x} + \frac{1}{x} C_1.$$

Уравнение Бернулли. Дифференциальное уравнение вида

$$y' + a(x) y = b(x) y^n, \quad n \neq 1,$$

называется *уравнением Бернулли*.

Решение уравнения Бернулли сводится к решению линейного уравнения делением обеих его частей на y^n и последующей заменой $z = \frac{1}{y^{n-1}}$:

$$\frac{1}{y^n} y' + a(x) \frac{1}{y^{n-1}} = b(x).$$

Далее, так как

$$z' = (1-n) \frac{1}{y^n} y',$$

то уравнение Бернулли сводится к линейному уравнению для неизвестной функции $z(x)$:

$$\frac{1}{1-n} z' + a(x) z = b(x).$$

6.4. Дифференциальные уравнения порядка выше первого. Примеры дифференциальных уравнений второго порядка. *Дифференциальным уравнением* n-го порядка, разрешенным относительно старшей производной, называется уравнение вида

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)}), \tag{8}$$

где x — независимая переменная, $y(x)$ — неизвестная функция, а f — заданная функция указанных аргументов.

Решением дифференциального уравнения (8) называется такая функция

$$y = y(x),$$

определенная на некотором промежутке $(a; b)$, которая при подстановке в уравнение (8) обращает уравнение в тождество для всех значений $x \in (a; b)$. Подстановка $y = y(x)$ в уравнение (8) возможна лишь в том случае, когда функция $y(x)$ имеет все производные до n-й включительно на промежутке $(a; b)$. Кроме того, также необходимо, чтобы при любом значении $x_0 \in (a; b)$ точка с координатами $(x_0; y(x_0); y'(x_0); \ldots; y^{(n-1)}(x_0))$ принадлежала области определения функции f.

Для дифференциального уравнения n-го порядка $y^{(n)} = f(x, y, y', \ldots, y^{(n-1)})$ имеет место

Теорема существования и единственности решения. *Пусть функция* $f(x, y, y', \ldots, y^{(n-1)})$ *определена в некоторой открытой области* D. *Если эта функция непрерывна в области* D *и все ее частные производные первого порядка по* $y, y', \ldots, y^{(n-1)}$ *также непрерывны в области* D, *то:*

1) *для любой точки* $(x_0; y_0; \dot{y}_0; \ldots; y_0^{(n-1)})$, *принадлежащей* D, *найдется функция* $y(x)$, *являющаяся решением дифференциального уравнения* (8) *и удовлетворяющая условиям*

$$y(x_0) = y_0, \quad y'(x_0) = \dot{y}_0, \quad \ldots, \quad y^{(n-1)}(x_0) = y_0^{(n-1)}; \tag{9}$$

2) *если* $y = \varphi(x)$ *и* $y = \psi(x)$ — *два решения уравнения* (8), *удовлетворяющие условиям* (9), *то эти два решения тождественно равны при всех значениях* x, *при которых они оба определены.*

Числа $x_0, y_0, \dot{y}_0, \ldots, y_0^{(n-1)}$ называются *начальными значениями* для решения $y(x)$, а равенство (9) — *начальными условиями.*

Простейшим дифференциальным уравнением порядка выше первого является дифференциальное уравнение второго порядка, разрешенное относительно второй производной

$$y'' = f(x, y, y').$$

Одним из примеров дифференциального уравнения второго порядка является дифференциальное уравнение *гармонических колебаний*

$$y'' + \omega^2 y = 0, \tag{10}$$

где ω — заданная постоянная, а y считается функцией времени t.

Для уравнения гармонических колебаний выполнены условия теоремы существования и единственности решения. Действительно, функция $f(t, y, y') = -\omega^2 y$ не зависит от t и y' и определена при любом y. Следовательно, область ее определения — все трехмерное пространство координат $(t; y, y')$. Далее, так как

$$\frac{\partial f}{\partial y} = -\omega^2 \quad \text{и} \quad \frac{\partial f}{\partial y'} = 0,$$

то частные производные функции f непрерывны при любых значениях t, y, y'.

Непосредственной подстановкой нетрудно проверить, что функция *)

$$y = A\cos(\omega t + \theta), \tag{11}$$

где A и θ — произвольные постоянные, является решением дифференциального уравнения (10). Постоянную A называют *амплитудой колебания*, ω — *частотой* колебания, а θ — *начальной фазой* колебаний.

Формула (11) дает множество всех решений дифференциального уравнения (10). Для выделения из бесконечного множества решений (11) единственного решения необходимо задать следующие начальные условия:

$$y(t_0) = y_0, \quad y'(t_0) = y_0',$$

т. е. задать значение функции $y(t)$ и ее производной $y'(t)$ в некоторый момент времени t_0, который называют *начальным моментом времени*. Задание значений y_0 и y_0' позволяет однозначно найти постоянные A и θ, выделяя тем самым из бесконечного множества решений (11) единственное.

Пример 1. Рассмотрим движение материальной точки массы m, скользящей по прямой l под действием переменной силы F, притягивающей материальную точку к точке O, лежащей на той же прямой. Пусть величина силы F пропорциональна величине отклонения материальной точки от точки O с некоторым коэффициентом пропорциональности k. (На практике такая сила F может быть осуществлена при помощи конструкции, состоящей из шара и двух пружин (рис. 9.27).) Здесь за точку O (начало координат) принимается середина отрезка BC. Отклонение материальной точки от середины будем обозначать через $x(t)$, где t — время. Ускорение материальной точки и сила, действующая на точку, связаны вторым законом Ньютона:

Рис. 9.27

$$mx'' = -kx, \tag{12}$$

где x'' — ускорение точки, $(-kx)$ — сила, действующая со стороны пружины на материальную точку, а $k > 0$ — коэффициент упругости пружины.

Согласно формуле (11) решение уравнения (12) имеет вид

$$x(t) = A\cos\left(\sqrt{\frac{k}{m}}\,t + \varphi_0\right),$$

откуда следует, что частота колебаний материальной точки

$$\omega = \sqrt{\frac{k}{m}}$$

определяется лишь ее массой и упругостью пружины k. Задавая начальные условия задачи, т. е. положение материальной точки в некоторый момент времени t_0 и ее скорость в начальный момент времени t_0:

$$x(t_0) \quad \text{и} \quad x'(t_0),$$

мы можем найти значения постоянных A и θ и тем самым полностью определить закон движения материальной точки.

Пример 2. Составим и решим приближенно уравнение математического маятника. Математический маятник представляет собой точку массы m, которая под действием силы тяжести движется по дуге окружности K радиуса l, лежащей в вертикальной плоскости (рис. 9.28). Величина l называется длиной маятника.

В плоскости окружности K введем полярную систему координат с центром в центре окружности и полярной осью, направленной вертикально вниз. Пере-

*) О методе нахождения этой функции см. п. 6 5, пример 1.

менный угол, характеризующий отклонение радиуса от вертикали, обозначим через $\varphi=\varphi(t)$. Точка M находится в поле силы тяжести $P=mg$, направленной вертикально вниз. Составляющая силы тяжести, направленная по радиусу окружности, уравновешивается реакцией связи (окружности, или нити, заставляющей точку двигаться по окружности); составляющая, направленная по касательной к окружности в точке M, равна $mg\sin\varphi$. Из физики известно, что линейное ускорение точки M связано с угловым ускорением φ'' формулой

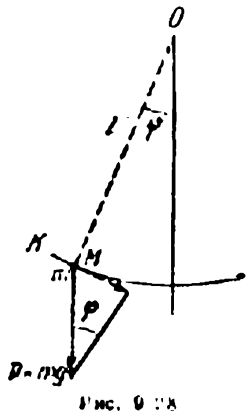


Рис. 9

$$a=l\varphi''.$$

Сила, действующая на точку, и ее ускорение связаны вторым законом Ньютона:

$$ml\varphi''=-mg\sin\varphi. \qquad (13)$$

Предполагая, что колебания, совершаемые материальной точкой, малы (т. е. угол отклонения $\varphi(t)$ мал), приближенно можно написать

$$\sin\varphi\approx\varphi,$$

и дифференциальное уравнение (13) принимает следующий вид

$$l\varphi''=-g\varphi. \qquad (14)$$

Решением дифференциального уравнения (14) является функция *)

$$\varphi(t)=A\cos\left(\sqrt{\frac{g}{l}}\,t+\varphi_0\right).$$

Из этой формулы следует, что частота «малых колебаний» математического маятника определяется только ускорением свободного падения g и длиной маятника l:

$$\omega=\sqrt{\frac{g}{l}}.$$

Для вычисления постоянных A и φ_0 необходимо задать положение и скорость точки M в некоторый момент времени t_0, после чего закон движения математического маятника становится полностью определенным.

6.5. Линейные дифференциальные уравнения с постоянными коэффициентами. *Линейным* дифференциальным уравнением n-го порядка с постоянными коэффициентами называется уравнение вида

$$a_0y^{(n)}+a_1y^{(n-1)}+\ldots+a_{n+1}y'+a_ny=f(x), \qquad (15)$$

где a_0, a_1, ..., a_{n-1}, a_n — действительные числа. Если $f(x)=0$, то уравнение (15) называется *однородным*, если же $f(x)\not\equiv 0$, то — *неоднородным*.

Общим решением линейного однородного уравнения n-го порядка называется функция

$$y=\varphi(x, C_1, C_2, \ldots, C_n),$$

зависящая от n произвольных постоянных C_1, C_2, ..., C_n и обращающая уравнение в тождество при любых значениях этих постоянных.

Решения линейного однородного уравнения

$$a_0y^{(n)}+a_1y^{(n-1)}+\ldots+a_{n-1}y'+a_ny=0 \qquad (16)$$

*) Подробнее о решении дифференциального уравнения (14), которое является неоднородным линейным уравнением второго порядка, см. п. 6.5.

будем искать в виде $y=e^{\lambda x}$. Подставляя $y=e^{\lambda x}$ в уравнение (16), получаем

$$(a_0\lambda^n+a_1\lambda^{n-1}+\ldots+a_{n-1}\lambda+a_n)\,e^{\lambda x}=0.$$

Поскольку $e^{\lambda x}\neq 0$, это уравнение эквивалентно алгебраическому уравнению

$$a_0\lambda^n+a_1\lambda^{n-1}+\ldots+a_{n-1}\lambda+a_n=0, \tag{17}$$

которое называют *характеристическим* уравнением.

Любая функция $y=e^{\lambda_i x}$, где λ_i — корень уравнения (17), является решением дифференциального уравнения (16).

Таким образом, чтобы найти решения линейного однородного дифференциального уравнения (16), необходимо написать характеристическое уравнение (17) и найти все его корни.

Общее решение линейного однородного дифференциального уравнения (16) представляется как сумма следующих функций:

1) $C_i e^{\lambda_i x}$, где λ_i — простой действительный корень уравнения (17) и C_i — произвольные постоянные;

2) $(C_{m+1}+C_{m+2}x+\ldots+C_{m+k}x^{k-1})\,e^{\lambda x}$, где λ — действительный корень кратности k, C_{m+1}, ..., C_{m+k} — произвольные постоянные;

3) $C_{l+1}e^{\alpha x}\cos\beta x+C_{l+2}e^{\alpha x}\sin\beta x$ для каждой пары простых комплексно сопряженных корней $\lambda=\alpha\pm i\beta$ уравнения (17); C_{l+1} и C_{l+2} — произвольные постоянные;

4) $P_{k-1}(x)\,e^{\alpha x}\cos\beta x+Q_{k-1}e^{\alpha x}\sin\beta x$ для каждой пары комплексно сопряженных корней кратности k. Здесь $P_{k-1}(x)$ и $Q_{k-1}(x)$ — многочлены $(k-1)$-й степени с произвольными постоянными коэффициентами.

Пример 1. Найти решение уравнения гармонических колебаний (см. п. 6.4, уравнение (10))

$$y''+\omega^2 y=0.$$

Характеристическое уравнение, соответствующее уравнению гармонических колебаний, имеет вид

$$\lambda^2+\omega^2=0.$$

Уравнение имеет два комплексно сопряженных корня

$$\lambda_1=i\omega,\quad \lambda_2=-i\omega,$$

и, следовательно, общее решение данного дифференциального уравнения имеет вид 3):

$$y=C_1\cos\omega t+C_2\sin\omega t,$$

где C_1 и C_2 — произвольные постоянные.

Методом дополнительного угла (см. п. 3.2 гл. 6) данную функцию можно записать в виде

$$y=A\sin(\omega t+\theta),$$

где A и θ — произвольные постоянные, выражающиеся через C_1 и C_2.

Пример 2. Решить уравнение

$$y^{(5)}-3y^{(4)}+y'''+y''+4y=0.$$

Пишем характеристическое уравнение

$$\lambda^5-3\lambda^4+\lambda^3+\lambda^2+4=0.$$

Корни характеристического уравнения найдем, используя способ, изложенный в п. 5.7 гл. 4;

$$\lambda_1=-1,\quad \lambda_2=\lambda_3=2,\quad \lambda_4=i,\quad \lambda_5=-i.$$

*) В соответствии с изложенными выше правилами общее решение уравнения (16) имеет вид

$$y=C_1e^{-x}+(C_2+C_3x)\,e^{2x}+C_4\cos x+C_5\sin x,$$

где C_1, C_2, C_3, C_4, C_5 — произвольные постоянные.

Частным решением линейного неоднородного уравнения называется любая конкретная функция, являющаяся решением этого уравнения.

Общее решение линейного неоднородного уравнения есть сумма общего решения линейного однородного уравнения, соответствующего неоднородному уравнению, и любого частного решения неоднородного уравнения. Приведем несколько простейших способов нахождения решений линейных неоднородных уравнений.

1. Если функция $f(x)$, стоящая в правой части уравнения, имеет вид

$$f(x)=P_m(x)\,e^{\gamma x},$$

где $P_m(x)$ — многочлен m-й степени, то частное решение линейного неоднородного уравнения ищется в виде

$$y=x^sQ_m(x)\,e^{\gamma x}, \tag{18}$$

где $Q_m(x)$ — многочлен той же степени m, что и многочлен $P_m(x)$. Число s равно нулю, если γ не является корнем характеристического уравнения; если же γ — корень характеристического уравнения, то s равно кратности этого корня. Чтобы найти коэффициенты многочлена $Q_m(x)$, необходимо подставить решение (18) в дифференциальное уравнение и приравнять коэффициенты при подобных членах в левой и правой частях уравнения.

Пример. Решить уравнение

$$y''-5y'+4y=4x^2e^{2x}. \tag{19}$$

Составляем характеристическое уравнение для данного уравнения и находим его корни:

$$\lambda^2-5\lambda+4=0,$$

$$\lambda_{1,2}=\frac{5\pm\sqrt{25-16}}{2};\quad \lambda_1=4,\quad \lambda_2=1.$$

Так как коэффициент, стоящий при x в показателе степени члена e^{2x}, равен 2 и не совпадает ни с одним из корней характеристического многочлена, то надо положить $s=0$, $x^0=1$ и частное решение искать в виде

$$y=(ax^2+bx+c)\,e^{2x}, \tag{20}$$

где коэффициенты a, b и c еще следует найти. Для этого подставим функцию, стоящую в правой части равенства (20), в уравнение (19). После приведения подобных членов в левой части уравнения получим

$$e^{2x}\left[-2ax^2+(-2a-2b)\,x+2a-b-2c\right]=4x^2e^{2x}.$$

Приравнивая коэффициенты при одинаковых степенях неизвестных, получаем систему для нахождения коэффициентов a, b и c:

$$\begin{aligned}-2a&=4;\\ -2a-2b&=0;\\ 2a-b-2c&=0,\end{aligned}$$

решение которой: $a=-2$, $b=2$, $c=-3$.

Следовательно, частное решение уравнения имеет вид

$$y=(-2x^2+2x-3)\,e^{2x}$$

и общее решение данного неоднородного уравнения следующее:

$$y=C_1e^{x}+C_2e^{4x}-(2x^2-2x+3)\,e^{2x},$$

где C_1 и C_2 — произвольные постоянные.

2. Если функция, стоящая в правой части уравнения, имеет вид

$$e^{\alpha x}(P(x)\cos\beta x+Q(x)\sin\beta x), \tag{21}$$

то частное решение уравнения ищется в виде

$$y=x^s e^{\alpha x}(M_k(x)\cos\beta x+N_k(x)\sin\beta x),$$

где $s=0$, если $\alpha+\beta i$ не является корнем характеристического многочлена; если же $\alpha+\beta i$ — корень характеристического многочлена, то s равно кратности этого корня; $M_k(x)$ и $N_k(x)$ — многочлены k-й степени, равной наибольшей из степеней многочленов $P(x)$ и $Q(x)$. Чтобы найти коэффициенты многочленов $M_k(x)$ и $N_k(x)$, необходимо подставить решение (21) в решаемое дифференциальное уравнение и приравнять коэффициенты при подобных членах в левой и правой частях уравнения.

Пример. Решить уравнение

$$y''+y=x\sin x. \tag{22}$$

Составляем характеристическое уравнение и находим его корни:

$$\lambda^2+1=0, \quad \lambda_1=i, \quad \lambda_2=-i.$$

Правая часть уравнения (22) — сокращенная запись выражения

$$e^{0\cdot x}(0\cdot\cos x+x\sin x).$$

Так как комплексное число $0+i$ является простым корнем характеристического уравнения, то частное решение уравнения (22) следует искать в виде

$$y=x(M(x)\cos x+N(x)\sin x),$$

где $M(x)$ и $N(x)$ — многочлены первой степени:

$$M(x)=ax+b, \quad N(x)=cx+d.$$

Подставляя функцию

$$x[(ax+b)\cos x+(cx+d)\sin x]$$

в уравнение (22) и приводя подобные члены в левой части уравнения, получаем

$$(4cx+2d+2a)\cos x+(-4ax+2c-2b)\sin x=x\sin x.$$

Приравнивая коэффициенты при подобных членах, стоящих в левой и правой частях последнего уравнения, получаем систему уравнений для нахождения коэффициентов a, b, c, d:

$$\begin{aligned} 4c&=0;\\ 2d+2a&=0;\\ -4a&=1;\\ 2c-2b&=0, \end{aligned}$$

решение которой: $a=-\frac{1}{4}$, $b=0$, $d=\frac{1}{4}$, $c=0$.

Следовательно, частное решение неоднородного уравнения имеет вид

$$y=-\frac{1}{4}x^2\cos x+\frac{1}{4}x\sin x,$$

а общее решение уравнения (22) —

$$y=C_1\sin x+C_2\cos x-\frac{1}{4}x^2\cos x+\frac{1}{4}x\sin x.$$

3. Если правая часть неоднородного уравнения равна сумме нескольких выражений вида (18) и (21), то частное решение линейного уравнения с правой частью вида $f_1+f_2+\ldots+f_k$ равно сумме частных решений уравнений с той же левой частью и правыми частями $f_1, f_2, \ldots, f_k$.

Метод вариации постоянных. Линейное неоднородное уравнение с постоянными коэффициентами

$$a_0y^{(n)}+a_1y^{(n-1)}+\ldots+a_ny=f(x) \tag{23}$$

с произвольной правой частью может быть решено методом, называемым методом *вариации постоянных*, который заключается в следующем.

Пусть $y=C_1y_1+C_2y_2+\ldots+C_ny_n$ — общее решение линейного однородного уравнения (23), где функции $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ находятся методом, изложенным выше. Тогда решение неоднородного уравнения (23) ищется в виде

$$y=C_1(x)\,y_1+C_2(x)\,y_2+\ldots+C_n(x)\,y_n,$$

где функции $C_1(x)$, $C_2(x)$, ..., $C_n(x)$ определяются из системы

$$\begin{aligned}
&C_1'y_1+C_2'y_2+\ldots+C_n'y_n=0;\\
&C_1'y_1'+C_2'y_2'+\ldots+C_n'y_n'=0;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&C_1'y_1^{(n-2)}+C_2'y_2^{(n-2)}+\ldots+C_n'y_n^{(n-2)}=0;\\
&C_1'y_1^{(n-1)}+C_2'y_2^{(n-1)}+\ldots+C_n'y_n^{(n-1)}=\frac{1}{a_0}f(x).
\end{aligned} \tag{24}$$

Пример. Решить уравнение

$$y''-y=2e^x-x^2 \tag{25}$$

методом вариации постоянных.

Составляем характеристическое уравнение и находим его корни:

$$\lambda^2-1=0;\quad \lambda_1=1,\quad \lambda_2=-1.$$

Общее решение однородного уравнения $y''-y=0$ имеет вид $y=C_1e^x+C_2e^{-x}$, где C_1 и C_2 — произвольные постоянные.

Общее решение уравнения (25) будем искать в виде

$$y=C_1(x)\,e^x+C_2(x)\,e^{-x}. \tag{26}$$

Составим для этого решения систему типа (24)

$$\begin{aligned}
&C_1'e^x+C_2'e^{-x}=0;\\
&C_1'e^x-C_2'e^{-x}=2e^x-x^2,
\end{aligned}$$

из которой получаем

$$\begin{aligned}
&C_1'=1-\frac{1}{2}x^2e^{-x};\\
&C_2'=\frac{1}{2}x^2e^x-e^{2x}.
\end{aligned}$$

Найдем функции $C_1(x)$ и $C_2(x)$:

$$C_1=\int\left(1-\frac{1}{2}x^2e^{-x}\right)dx=x+\frac{1}{2}x^2e^{-x}+xe^{-x}+e^{-x}+C_1^0;$$

$$C_2=\int\left(\frac{1}{2}x^2e^x-e^{2x}\right)dx=\frac{1}{2}x^2e^x-xe^x+e^x-\frac{1}{2}e^{2x}+C_2^0,$$

где C_1^0 и C_2^0 — постоянные интегрирования.

Подставляя найденные функции $C_1(x)$ и $C_2(x)$ в формулу (26), получаем общее решение уравнения (25)

$$y=\left(C_1^0-\frac{1}{2}\right)e^x+C_2^0e^{-x}+xe^x+x^2+2.$$

Обозначая $C_1^0-\frac{1}{2}=C_1$ и $C_2^0=C_2$, окончательно имеем

$$y=C_1e^x+C_2e^{-x}+xe^x+x^2+2.$$

Глава 10

КОМБИНАТОРИКА. ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Комбинаторика — раздел математики о выборе и расположении элементов некоторого множества на основании каких-либо условий.

Комбинаторика начала выделяться в отдельный раздел математики в работах Б. Паскаля и П. Ферма, хотя отдельные понятия и факты комбинаторики были известны еще математикам античности и средневековья. Большой вклад в развитие комбинаторики внесли Г. Лейбниц, Я. Бернулли и Л. Эйлер. В их работах были даны определения основных понятий комбинаторики, развиты первые комбинаторные методы и указаны их применения, а также прослежена связь комбинаторики с исчислением вероятностей. Именно комбинаторика послужила фундаментальной основой началам теории вероятностей.

Теория вероятностей — раздел математики, посвященный изучению вероятности — меры возможности наступления какого-либо определенного события в тех или иных определенных условиях, которые могут повторяться сколь угодно большое число раз. Теория вероятностей позволяет по вероятностям наступления одних случайных событий находить вероятность наступления других случайных событий, каким-либо образом связанных с первыми.

§ 1. Метод математической индукции

Ряд математических утверждений доказывается методом математической индукции, основанном на *принципе математической индукции*, который заключается в следующем.

Пусть $A(n)$ — некоторое утверждение, имеющее смысл для натуральных чисел n, и пусть это утверждение истинно для $n=1$. Тогда если из истинности этого утверждения для $n=k$ (k — любое натуральное число, большее единицы) следует истинность утверждения для $n=k+1$, то утверждение $A(n)$ истинно для любого натурального числа n.

Доказательство *методом математической индукции* состоит в следующем:

1) проверяется, что утверждение $A(n)$ истинно для $n=1$;

2) предполагается, что утверждение $A(n)$ истинно для $n=k$ и доказывается, что оно истинно и при $n=k+1$.

После этого на основании принципа математической индукции делается вывод, что утверждение $A(n)$ истинно для любого натурального n.

Пример. Доказать, что для любого натурального числа n и любого действительного числа $a > -1$ справедливо неравенство, называемое *неравенством Бернулли*:

$$(1+a)^n \geq 1+an. \qquad (1)$$

Если $n=1$, то справедливость неравенства (1) очевидна:

$$(1+a)^1 \geqslant 1+a.$$

Предположим, что неравенство (1) справедливо при $n=k$:

$$(1+a)^k \geqslant 1+ak.$$

Умножая обе части последнего неравенства на положительное число $1+a$, получаем неравенство

$$(1+a)^{k+1} \geqslant 1+ak+a+a^2k. \qquad (2)$$

Отбрасывая последнее положительное слагаемое в правой части неравенства (2), мы уменьшаем правую часть этого неравенства, а потому

$$(1+a)^{k+1} \geqslant 1+a(k+1).$$

Полученный результат показывает, что из предположения о справедливости неравенства (1) при $n=k$ следует, что неравенство (1) справедливо и при $n=k+1$.

Обе части доказательства методом математической индукции проведены, и, следовательно, неравенство (1) справедливо при любом натуральном n.

Методом математической индукции может быть доказан также, например, ряд свойств чисел, образующих последовательность Фибоначчи. *Последовательностью Фибоначчи* называется последовательность чисел $F_0, F_1, F_2, \ldots, F_n, \ldots$, где $F_0=0$, $F_1=1$, а каждый последующий член является суммой двух предыдущих, т. е. указанная последовательность имеет вид 0, 1, 1, 2, 3, 5, 8, ...

Приведенная последовательность дает решение задачи, впервые предложенной в 1202 г. Леонардо Фибоначчи, которая формулировалась так: «Сколько пар кроликов получится от одной пары в течение года? При этом предлагается принять, что каждая пара дает приплод — пару кроликов — каждый месяц, каждая новая пара становится плодоносящей в возрасте одного месяца и, кроме того, кролики никогда не мрут».

Докажем, что если через Φ *) обозначить число $(1+\sqrt{5})/2$, то

$$F_n < \Phi^{n-1} \qquad (3)$$

для всех натуральных n.

Если $n=1$, то $F_1=1=\Phi^0$, т. е. нестрогое неравенство (3) выполнено.

Предположим, что для любого натурального $k>1$ выполняются неравенства

$$F_2 \leqslant \Phi^1, \quad F_3 < \Phi^2, \ldots, \quad F_k \leqslant \Phi^{k-1}.$$

Докажем, что тогда неравенство (3) справедливо и при $n=k+1$. Действительно, $F_{k+1}=F_{k-1}+F_k \leqslant \Phi^{k-2}+\Phi^{k-1}=\Phi^{k-2}(1+\Phi)$. Замечая, что $1+\Phi=\Phi^2$, окончательно получаем

$$F_{k+1} \leqslant \Phi^{k-2}\Phi^2=\Phi^k.$$

Таким образом, из предположения о справедливости неравенства (3) при $n=k$ следует его справедливость и при $n=k+1$.

*) Число Φ называют *отношением крайнего и среднего*, так как оно определяется условием: если $\Phi=A:B$, то $A:B=(A+B):A$. Известны и другие названия этого числа: *божественная пропорция*, *золотое сечение*. Само обозначение Φ происходит от первой буквы имени греческого скульптора Фидия, который часто использовал золотое сечение в пропорциях своих скульптур.

Обе части доказательства методом математической индукции проведены, и, следовательно, высказанное утверждение справедливо при любом натуральном n.

Методом математической индукции также можно доказать справедливость для всех натуральных n следующих соотношений:

1) $F_{n+1} \geqslant \Phi^{n-1}$;

2) $F_{n+1} \cdot F_{n-1} - F_n^2 = (-1)^n$;

3) $F_{n+m} = F_m \cdot F_{n+1} + F_{m-1} \cdot F_n$ для любого натурального m.

В ряде случаев может оказаться, что рассматриваемое утверждение $A(n)$ при некоторых натуральных значениях $n < m$ (m — фиксированное натуральное число) ложно или не имеет смысла. С другой стороны, в ряде случаев оказывается возможным доказать, что некоторое утверждение $A(n)$ истинно не только для натуральных значений n, но и для некоторого набора отрицательных значений n. В этих случаях используется следующее обобщение принципа математической индукции.

Если утверждение $A(n)$, в котором n — целое число, истинно при $n = m$ и если из истинности этого утверждения для числа $n = k$, где k — любое целое число, большее или равное m, вытекает, что оно истинно для следующего числа $n = k + 1$, то утверждение $A(n)$ истинно и для любого целого числа $n \geqslant m$.

§ 2. Комбинаторика

2.1. Перестановки. Пусть A — некоторое конечное множество, состоящее из n различных элементов:

$$A = \{a_1;\ a_2;\ a_3;\ \ldots;\ a_n\}. \tag{1}$$

Будем образовывать из элементов множества A упорядоченные множества. В качестве первого упорядоченного множества возьмем множество, в котором элементы расположены в порядке возрастания их номеров:

$$(a_1;\ a_2;\ a_3;\ \ldots;\ a_n).$$

Второе упорядоченное множество можно образовать, поменяв местами элементы a_1 и a_2, а все остальные элементы первого множества оставив на своих местах:

$$(a_2;\ a_1;\ a_3;\ \ldots;\ a_n).$$

Поменяв местами элементы a_1 и a_3 и оставляя на своих местах все остальные элементы в первом упорядоченном множестве, получаем упорядоченное множество, отличное как от первого, так и от второго упорядоченного множества. Аналогичным способом из элементов множества A можно строить и другие упорядоченные множества.

Всевозможные конечные упорядоченные множества, содержащие n различных элементов, которые можно получить из некоторого неупорядоченного множества, состоящего из n различных элементов, называются *перестановками* из n элементов.

Таким образом, перестановка есть не что иное, как способ упорядочивания элементов некоторого конечного множества. При этом любые две различные перестановки представляют собой два различных способа образования упорядоченного множества (из данного неупорядоченного множества).

Вычислим, сколько различных упорядоченных множеств можно образовать из данного конечного множества A, содержащего n различных элементов, т. е. чему равно число перестановок из элементов множества (1).

Любое упорядоченное множество, полученное из множества A, можно записать в виде

$$(i_1;\ i_2;\ \ldots;\ i_k;\ \ldots;\ i_n),$$

где каждый из элементов i_k есть один из элементов $a_1, a_2, a_3, \ldots, a_n$, причем ни один из этих элементов не встречается более одного раза. В качестве элемента i_1 можно взять любой из элементов a_1, a_2, a_3, $\ldots$, a_n, что дает n различных возможностей. Теперь, если i_1 уже выбран, в качестве i_2 можно взять любой из $n-1$ оставшихся элементов множества A, т. е. число различных способов выбора упорядоченной пары элементов $(i_1; i_2)$ будет равно $(n-1)\cdot n$. Продолжая далее указанный процесс, в результате получим, что последний оставшийся элемент i_n можно выбрать лишь одним способом.

Таким образом, число перестановок из n элементов (которое обычно обозначается P_n) равно произведению n последовательных натуральных чисел, начиная с единицы. Это произведение имеет специальное обозначение $n!$ (читается: n *факториал*):

$$P_n = 1\cdot 2\cdot 3\ \ldots\ (n-1)\cdot n = n!$$

Для пустого множества принимается соглашение: пустое множество можно упорядочить только одним способом; поэтому считается, что

$$0! = 1.$$

Если в заданной перестановке поменять местами какие-либо два элемента, а все остальные элементы оставить на своих местах, то получится новая перестановка.

Такое преобразование перестановки называется *транспозицией*.

Все $n!$ перестановок из n элементов можно расположить в таком порядке, что каждая следующая перестановка будет получаться из предыдущей одной транспозицией, причем в качестве исходной перестановки можно выбрать любую из $n!$ перестановок. В частности, от любой перестановки из n элементов можно перейти к любой другой перестановке из тех же элементов при помощи нескольких транспозиций.

Перестановка $(i_1; i_2; \ldots; i_n)$ называется *четной*, если она получается из перестановки $(a_1; a_2; \ldots; a_n)$ с помощью четного числа транспозиций, и *нечетной* в противном случае.

Пример. Пусть $A = (1; 2; 3; 4)$. Перестановка (1; 2; 3; 4) по определению четна. Четной будет и перестановка (4; 2; 1; 3), так как она может быть получена из перестановки (1; 2; 3; 4) при помощи четного числа транспозиций: поменяем местами в перестановке (4; 2; 1; 3) первый и четвертый элементы:

$$\overline{(4;\ 2;\ 1;\ 3)} \to (3;\ 2;\ 1;\ 4);$$

в полученной перестановке меняем местами первый и третий элементы:

$$\overleftarrow{(3;\ 2;\ 1;\ 4)} \to (1;\ 2;\ 3;\ 4).$$

Таким образом, данная перестановка переходит в исходную в результате двух транспозиций, и, следовательно, она четная. Аналогичным образом можно убе-

дится, что перестановка (4; 1; 2; 3) будет нечетной — она переходит в перестановку (1; 2; 3; 4) после трех транспозиций:

$$(\overleftarrow{4;\ 1};\ 2;\ 3) \to (1;\ \overleftarrow{4;\ 2};\ 3) \to (1;\ 2;\ \overleftarrow{4;\ 3}) \to (1;\ 2;\ 3;\ 4).$$

Всякая транспозиция меняет четность перестановки: число четных перестановок из n элементов ($n \geqslant 2$) равно числу нечетных, т. е. равно $n!/2$.

2.2. Подстановки. Пусть A — некоторое конечное множество, состоящее из n различных элементов. Для простоты будем считать, что это первые n натуральных чисел, т. е.

$$A = (1;\ 2;\ 3;\ \ldots;\ n).$$

Всякое взаимно однозначное отображение F множества A на себя называется *подстановкой* n-й степени.

Для обозначения подстановки F используется запись

$$F = \begin{pmatrix} i_1 & i_2 & i_3 & \ldots & i_n \\ k_1 & k_2 & k_3 & \ldots & k_n \end{pmatrix},$$

где $(i_1;\ i_2;\ i_3;\ \ldots;\ i_n)$, $(k_1;\ k_2;\ k_3;\ \ldots;\ k_n)$ — две перестановки из элементов множества A ($k_1, k_2, k_3, \ldots, k_n$ — образы элементов $i_1, i_2, i_3, \ldots, i_n$ при отображении F).

Одна и та же подстановка может быть записана многими способами. Так, подстановка (2) также может быть записана в виде

$$\begin{pmatrix} i_2 & i_1 & i_3 & \ldots & i_n \\ k_2 & k_1 & k_3 & \ldots & k_n \end{pmatrix} \quad \text{или} \quad \begin{pmatrix} i_n & i_{n-1} & \ldots & i_3 & i_2 & i_1 \\ k_n & k_{n-1} & \ldots & k_3 & k_2 & k_1 \end{pmatrix}.$$

В частности, всякая подстановка n-й степени может быть записана в виде

$$F = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ l_1 & l_2 & l_3 & \ldots & l_n \end{pmatrix},$$

т. е. в порядке возрастания натуральных чисел в верхней строке. При такой записи различные подстановки отличаются друг от друга перестановками, стоящими в нижней строке, а потому число подстановок n-й степени равно числу перестановок из n элементов, т. е. равно $n!$

Тождественной подстановкой n-й степени называется подстановка

$$E = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 1 & 2 & 3 & \ldots & n \end{pmatrix}.$$

Рассмотрим теперь произвольную подстановку n-й степени

$$F = \begin{pmatrix} i_1 & i_2 & i_3 & \ldots & i_n \\ k_1 & k_2 & k_3 & \ldots & k_n \end{pmatrix}.$$

Перестановки из n элементов, образующие верхнюю и нижнюю строки подстановки F, имеют либо одинаковые четности (т. е. они или обе четны, или обе нечетны), либо противоположные четности (т. е. одна четная, а другая нечетная), причем совпадение или несовпадение четностей перестановок не зависит от различных способов записи данной подстановки.

Если четности верхней и нижней строк подстановки F совпадают, то подстановка называется *четной*; если они противоположны — то *нечетной*.

Число четных подстановок n-й степени равно числу нечетных, т. е. равно $n!/2$.

Умножение подстановок. По определению подстановка n-й степени есть взаимно однозначное отображение F конечного множества из n различных элементов на себя. Последовательное выполнение двух подстановок n-й степени опять будет взаимно однозначным отображением множества A на себя и приводит к некоторой вполне определенной третьей подстановке n-й степени, называемой *произведением* первой из заданных подстановок на вторую.

О произведении двух данных подстановок F и G также говорят, что оно получено в результате умножения подстановки F на подстановку G. Произведение подстановки F на подстановку G обозначают FG.

Например, пусть даны две подстановки четвертой степени

$$F = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 2 & 3 & 1 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}.$$

Произведением подстановок F и G является подстановка

$$FG = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}.$$

Действительно, при подстановке F число 1 переходит в 4, а при подстановке G число 4 переходит в 2, поэтому при отображении FG число 1 переходит в 2 и т. д.

Умножение подстановок определено лишь для подстановок одинаковой степени и обладает следующими свойствами:

1) умножение подстановок n-й степени при $n \geqslant 3$ некоммутативно, т. е. подстановка FG не всегда равна подстановке GF;

2) умножение подстановок ассоциативно: $F(GH) = (FG)H$; поэтому имеет смысл говорить о произведении любого конечного числа подстановок n-й степени;

3) произведение любой подстановки F на тождественную подстановку E, а также произведение E на F равны F:

$$EF = FE = F.$$

Обратной подстановкой к подстановке n-й степени F называется такая подстановка n-й степени G, что

$$FG = GF = E.$$

Подстановка, обратная подстановке F, обычно обозначается F^{-1}:

$$FF^{-1} = F^{-1}F = E.$$

Подстановка, обратная подстановке

$$F = \begin{pmatrix} i_1 & i_2 & i_3 & \dots & i_n \\ k_1 & k_2 & k_3 & \dots & k_n \end{pmatrix},$$

имеет вид

$$F^{-1} = \begin{pmatrix} k_1 & k_2 & k_3 & \dots & k_n \\ i_1 & i_2 & i_3 & \dots & i_n \end{pmatrix},$$

т. е. получается из F перестановкой строк.

2.3. Размещения. Пусть дано некоторое конечное множество A, состоящее из n различных элементов. Выберем некоторым образом из n элементов m различных элементов и будем составлять из этих m элементов различные упорядоченные множества.

Конечные упорядоченные подмножества, содержащие по m элементов, выбранных из n элементов основного множества, называются *размещениями* из n элементов по m элементов. Число всех возможных размещений из n элементов по m обозначается A_n^m.

Нетрудно убедиться, что

$$A_n^1 = n,$$

т. е. один элемент из n можно выбрать n способами, а из одного элемента можно образовать единственное упорядоченное множество.

Найдем, чему равно число размещений из n элементов по m. Чтобы распределить $m+1$ элементов, взятых из n элементов, по $m+1$ местам, можно сначала выбрать m элементов и разместить их по первым m местам. Это можно сделать A_n^m способами. При каждом таком выборе m элементов из n остаются «свободными» $n-m$ элементов, любой из которых можно поставить на свободное $(m+1)$-е место. Таким образом, при каждом из A_n^m заполнений первых m мест получим $n-m$ возможных заполнений $(m+1)$-го места. Следовательно, число размещений по $m+1$ элементов, выбранных из n заданных элементов, связано с числом размещений из n элементов по m равенством

$$A_n^{m+1} = (n-m)\,A_n^m.$$

Учитывая, что $A_n^1 = n$, имеем

$$A_n^1 = n,$$
$$A_n^2 = n(n-1),$$
$$A_n^3 = n(n-1)(n-2),$$
$$\cdots\cdots\cdots$$
$$A_n^m = n(n-1)(n-2)\ldots(n-m+1).$$

Последнее равенство можно записать с использованием символа факториала:

$$A_n^m = \frac{n!}{(n-m)!}.$$

При $m=0$ получаем $A_n^0 = 1$, т. е. из любого множества единственным способом можно извлечь пустое множество, которое, по определению, можно упорядочить единственным способом.

Число размещений и число перестановок связаны формулой

$$A_n^n = P_n = n!$$

2.4. Сочетания. Конечные неупорядоченные множества, содержащие m различных элементов, выбранных из n элементов заданного множества, называются *сочетаниями* из n элементов по m. Число сочетаний из n элементов по m элементов обозначается C_n^m или $\binom{n}{m}$.

Найдем, чему равно число сочетаний из n элементов по m. В согласии с данным выше определением размещений, из данного множества, содержащего n элементов, можно образовать A_n^m различных упорядоченных множеств, содержащих по m различных элементов. Из множества, содержащего m различных элементов, можно образовать P_m различных упорядоченных множеств, а потому число C_n^m различных неупорядоченных множеств, содержащих по m раз-

личных элементов, выбранных из элементов, будет вычисляться по формуле

$$C_n^m = \frac{A_n^m}{P_m}.$$

Используя формулы для подсчета числа перестановок P_m и числа размещений A_n^m, получим

$$C_n^m = \frac{n!}{m!\,(n-m)!}.$$

Для числа сочетаний справедливы равенства

$$C_n^m = C_n^{n-m}, \quad C_{n+1}^{m+1} = C_n^m + C_n^{m+1}, \tag{3}$$

а также

$$C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^{n-1} + C_n^n = 2^n.$$

Последнее свойство иногда формулируется в виде следующей теоремы о конечных множествах.

Число всех подмножеств множества, состоящего из n элементов, равно 2^n.

В табл. 9 даны значения C_n^m для $n = 10$ и $m \leqslant n$.

Эта таблица называется *треугольником Паскаля*. Отметим, что вторая из формул (3) позволяет и далее последовательно заполнять

Таблица 9

n	m										
	0	1	2	3	4	5	6	7	8	9	10
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1

строки треугольника Паскаля, пользуясь тем, что в начале и в конце каждой строки стоят единицы. Вычислим, например, C_{11}^4, не приведенное в таблице:

$$C_{11}^4 = C_{10}^3 + C_{10}^4 = 120 + 210 = 330.$$

2.5. Бином Ньютона. Натуральная степень суммы двух величин вычисляется по формуле

$$(a+b)^n = C_n^0 a^n + C_n^1 a^{n-1} b + \ldots + C_n^m a^{n-m} b^m + \ldots + C_n^n b^n = \sum_{i=0}^{n} C_n^i a^{n-i} b^i,$$

которая называется *биномом Ньютона*. Коэффициенты C_n^m, равные числу сочетаний из n элементов по m, называются *биномиальными коэффициентами*.

Свойства биномиальных коэффициентов:

1) число биномиальных коэффициентов (а следовательно, и число слагаемых в разложении степени бинома) равно $n+1$;

2) коэффициенты членов, равноотстоящих от начала и от конца бинома, равны между собой (так как $C_n^m = C_n^{n-m}$);

3) сумма всех биномиальных коэффициентов равна 2^n;

4) сумма биномиальных коэффициентов, стоящих на нечетных местах, равна сумме коэффициентов, стоящих на четных местах.

Натуральная степень разности двух величин вычисляется по формуле, аналогичной биному Ньютона:

$$(a-b)^n = C_n^0 a^n - C_n^1 a^{n-1}b + \ldots + (-1)^m C_n^m a^{n-m}b^m + \ldots + (-1)^n C_n^n b^n = \sum_{i=0}^{n} (-1)^i C_n^i a^{n-i}b^i.$$

2.6. Перестановки с повторениями. Пусть A — некоторая совокупность, состоящая из n элементов m различных типов ($m \leqslant n$)

$$A = \{a_1, a_2, a_3, \ldots, a_n\},$$

причем элементы одного типа неразличимы между собой. И пусть k_1 элементов принадлежат первому типу, k_2 — второму, k_3 — третьему, k_m — m-му типу, причем $k_1 + k_2 + k_3 + \ldots + k_m = n$.

Например, если

$$A = \{1, 1, 2, 3, 1, 3\}, \tag{4}$$

то, считая элементами первого типа единицы, элементами второго типа двойки, а элементами третьего типа тройки, имеем $k_1 = 3$, $k_2 = 1$, $k_3 = 2$.

Различные конечные совокупности, содержащие n элементов, из которых k_1 принадлежат первому типу, k_2 — второму, ..., k_m — m-му типу, называются *перестановками с повторениями* (*кортежами*), *имеющими состав* (или *спецификацию*) $(k_1, k_2, k_3, \ldots, k_m)$.

Число таких различных перестановок с повторениями обычно обозначается $C_n(k_1, k_2, \ldots, k_m)$ и вычисляется по формуле

$$C_n(k_1, k_2, \ldots, k_m) = \frac{n!}{k_1! k_2! \ldots k_m!}. \tag{5}$$

Формула (5) может быть доказана следующим образом. Рассмотрим некоторую конкретную перестановку с повторениями состава $(k_1, k_2, \ldots, k_m)$ и заменим все одинаковые элементы разными. В результате такой замены получим некоторую перестановку из n различных элементов. Число способов, которыми можно заменить k_1 неразличимых элементов первого типа разными элементами, равно $k_1!$; число способов, которыми можно заменить k_2 неразличимых элементов второго типа разными, равно $k_2!$ и т. д. И, наконец, число способов, которыми можно заменить k_m неразличимых элементов разными, равно $k_m!$

Таким образом, число различных перестановок из n различных элементов, которые могут быть получены из одной конкретной перестановки с повторениями состава $(k_1, k_2, \ldots, k_m)$, равно $k_1! \cdot k_2! \cdot \ldots \cdot k_m!$

С другой стороны, число перестановок из n различных элементов равно $n!$, и поэтому число всех перестановок с повторениями состава $(k_1, k_2, \ldots, k_m)$ связано с числом перестановок из n различных элементов равенством

$$C_n(k_1, k_2, \ldots, k_m) \cdot k_1! \cdot k_2! \cdot \ldots \cdot k_m! = n!,$$

из которого и следует формула (5).

Пример. Сколько различных шестизначных чисел можно образовать из трех единиц, одной двойки и двух троек?

Число таких различных шестизначных чисел равно числу различных перестановок с повторениями состава (3, 1, 2)

$$C_6(3, 1, 2) = \frac{6!}{3! \cdot 1! \cdot 2!} = 60.$$

Число перестановок с повторениями состава $(k_1, k_2, \ldots, k_m)$ может быть введено как решение задачи о числе разбиений множества, состоящего из n различных элементов, на сумму m непересекающихся подмножеств.

Пусть A — некоторое конечное множество, состоящее из n различных элементов:

$$A = \{a_1; a_2; a_3; \ldots; a_n\}.$$

Представим множество A в виде суммы m подмножеств $B_1, B_2, \ldots, B_m$, не имеющих общих элементов:

$$A = B_1 \cup B_2 \cup B_3 \cup \ldots \cup B_m.$$

Пусть число элементов в подмножестве B_1 равно k_1, число элементов в подмножестве B_2 равно k_2, ..., число элементов в подмножестве B_m равно k_m $(k_1 + k_2 + \ldots + k_m = n)$. Подсчитаем число всех возможных разбиений множества A на сумму фиксированного числа подмножеств $B_1, B_2, \ldots, B_m$.

Все разбиения множества A на m групп $B_1, B_2, \ldots, B_m$ можно получить следующим образом.

Образуем подмножество B_1 из k_1 элементов множества A. Это можно сделать $C_n^{k_1}$ способами. Из оставшихся $n - k_1$ элементов множества A образуем множество B_2, содержащее k_2 элементов. Это можно сделать $C_{n-k_1}^{k_2}$ способами. Общее число способов образования двух таких подмножеств B_1 и B_2 равно

$$C_n^{k_1} \cdot C_{n-k_1}^{k_2} = \frac{n!}{k_1!\,(n-k_1)!} \cdot \frac{(n-k_1)!}{k_2!\,(n-k_1-k_2)!} = \frac{n!}{k_1!\,k_2!\,(n-k_1-k_2)!}.$$

Продолжая далее процедуру образования подмножеств $B_3, \ldots, B_m$, в результате получим, что общее число способов выбора подмножеств $B_1, B_2, \ldots, B_m$ равно

$$C_n^{k_1} \cdot C_{n-k_1}^{k_2} \cdot \ldots \cdot C_{n-k_1-\ldots-k_{m-1}}^{k_m} = \frac{n!}{k_1! \cdot k_2! \cdot \ldots \cdot k_m!}.$$

Таким образом, число способов, которыми можно представить множество A из n различных элементов в виде суммы m подмножеств $B_1, B_2, \ldots, B_m$, число элементов которых равно $k_1, k_2, \ldots, k_m$ соответственно $(k_1 + k_2 + \ldots + k_m = n)$, равно

$$\frac{n!}{k_1!\,k_2!\,\ldots\,k_m!}$$

и является числом различных перестановок с повторениями $C_n(k_1, k_2, \ldots, k_m)$.

Числа $C_n(k_1, k_2, \ldots, k_m)$ также называют *полиномиальными коэффициентами* — эти числа являются коэффициентами в формуле вычисления n-й степени суммы m величин

$$(a_1+a_2+\ldots+a_m)^n=\sum C_n(k_1, k_2, \ldots, k_m)\, a_1^{k_1}\cdot a_2^{k_2}\cdot\ldots\cdot a_m^{k_m}. \qquad (6)$$

В правой части равенства (6) стоит сумма всевозможных слагаемых вида $C_n(k_1, k_2, \ldots, k_m)\, a_1^{k_1} a_2^{k_2}\ldots a_m^{k_m}$, где целые числа k_1, k_2, $\ldots$, k_m связаны равенством $k_1+k_2+\ldots+k_m=n$.

Если $a_1=a_2=\ldots=a_m$, то из формулы (6) следует равенство

$$\sum C_n(k_1, k_2, \ldots, k_m)=m^n.$$

При $m=2$ формула (6) представляет собой формулу бинома Ньютона, полиномиальные коэффициенты превращаются в биномиальные и их сумма равна 2^m.

2.7. **Сочетания с повторениями.** *Сочетаниями из m (различных) элементов по n элементов с повторениями* называют неупорядоченные совокупности, состоящие из n элементов, каждый из которых принадлежит к одному из m типов.

Например, из трех различных элементов a_1, a_2, a_3 можно составить следующие сочетания по два с повторениями:

$$\{a_1; a_1\},\ \{a_1; a_2\},\ \{a_1; a_3\},$$
$$\{a_2; a_2\},\ \{a_2; a_3\},\ \{a_3; a_3\}.$$

Число различных сочетаний из m элементов по n с повторениями обозначается f_m^n и вычисляется по формуле

$$f_m^n=C_{m+n-1}^{n}=C_{m+n-1}^{m-1}. \qquad (7)$$

Справедливость этой формулы нетрудно установить на основании следующих рассуждений. Каждое сочетание из m элементов по n с повторениями полностью определяется, если указать, сколько элементов каждого из m типов в него входит. Поставим в соответствие каждому сочетанию с повторениями последовательность нулей и единиц, составленную по следующему правилу: напишем подряд столько единиц, сколько элементов первого типа входит в сочетание (если элемент первого типа не входит в сочетание — не пишем ничего); после этого поставим нуль, обозначая им «границу» между элементами первого и второго типов. Напишем подряд столько единиц, сколько элементов второго типа входит в сочетание (если элемент второго типа вообще не входит в сочетание, то после нуля не пишем ничего); ставим нуль, указывая, что количество элементов второго типа перечислено, и т. д. Например, написанным выше сочетаниям из трех элементов по два с повторениями будут соответствовать следующие последовательности нулей и единиц:

1100, 1010, 1001, 0110, 0101, 0011.

Таким образом, каждому сочетанию из m по n с повторениями можно поставить в соответствие последовательность из n единиц и $m-1$ нулей (число «границ» между m типами элементов), и наоборот, каждой такой последовательности единиц и нулей соответствует единственное сочетание.

Число способов, которыми можно расставить n единиц по $n+m-1$ местам, равно C_{n+m-1}^{n}. С другой стороны, число способов,

которыми можно расставить $m-1$ нулей по $n+m-1$ местам, равно C_{n+m-1}^{m-1}. По свойству 2) биномиальных коэффициентов число сочетаний $C_{n+m-1}^{m-1}=C_{n+m-1}^{n}$, и формула (7) доказана.

Для числа сочетаний с повторениями справедлива следующая рекуррентная формула:

$$\overline{C}_n^m=\overline{C}_{n-1}^m+\overline{C}_n^{m-1}.$$

§ 3. Элементы теории вероятностей

3.1. Испытание. Исход. Событие. В теории вероятностей под *испытанием* (*опытом*) понимают комплекс определенных условий, воспроизводимый сколь угодно большое число раз, наличие которого приводит к какому-либо *исходу*. Исход случаен, если испытание, приводящее к какому-либо определенному исходу, имеет и другие исходы.

Случайным событием (или просто *событием*) называется любая совокупность случайных исходов. Событие может состоять и из одного исхода, и поэтому сами исходы также можно именовать событиями. Исходы называются *независимыми*, если они представляют собой результаты независимых испытаний.

Пример. Возьмем игральную кость (кубик, на гранях которого отмечены от одного до шести очков) и бросим его на стол. На верхней грани брошенного кубика выпадет какое-либо количество очков. Совокупность указанных условий (наличие у нас кубика, стола и процедура бросания кубика) — испытание. Количество очков, выпавших на верхней грани брошенного кубика, — исход этого испытания. Общее число случайных исходов равно шести. Случайным событием в данном примере будет, например, событие, состоящее в том, что на верхней грани окажется четное число очков. Это событие есть совокупность трех исходов: выпадения 2, 4 и 6 очков.

Случайные события обычно обозначают большими латинскими буквами A, B, C, ... или A_1, A_2, A_3, ...

Достоверным событием называют событие, которое обязательно произойдет при данном испытании. *Невозможным* событием называют событие, которое заведомо не произойдет при данном испытании. Так, в рассмотренном испытании с игральной костью событие, состоящее в том, что в результате испытания выпало некоторое число очков, не большее шести, — достоверное, а событие, состоящее в том, что выпало семь очков, — невозможное.

Два случайных события называются *противоположными*, если одно из них происходит в том и только в том случае, когда не происходит другое. Событие, противоположное событию A, обычно обозначают $\overline{A}$.

Пример. Брошена игральная кость. Обозначим через A событие, которое состоит в том, что выпало шесть очков. Событие $\overline{A}$ — выпадение от одного до пяти очков.

События A_1, A_2, ..., A_n называются *несовместными*, если наступление одного из них исключает наступление других.

Пример. Пусть в урне находятся три белых шара, занумерованных цифрами 1, 2, 3, и два черных шара, занумерованных цифрами 4, 5. Вытаскиваем наугад один из шаров. Пусть событие A состоит в том, что вытащенный шар — белый, а событие B состоит

в том, что вытащенный шар имеет четный номер. События A и B — совместные, так как из урны может быть вынут белый шар с номером 2. Далее, пусть событие C состоит в том, что вынут шар под номером 5. События A и C несовместны, так как белые шары занумерованы цифрами 1, 2, 3.

Говорят, что события $A_1, A_2, \ldots, A_n$ образуют *полную группу событий*, если в результате данного испытания обязательно произойдет одно из них.

Пример. Пусть в урне находятся три белых шара, занумерованных цифрами 1, 2, 3, и два черных шара, занумерованных цифрами 1, 2. Из урны наугад вынимают один шар. События:

A_1 — появление шара с цифрой 1,
A_2 — появление шара с цифрой 2,
A_3 — появление шара с цифрой 3

образуют полную группу событий.

Полная группа событий $A_1, A_2, \ldots, A_n$ называется *полной группой несовместных событий*, если в результате данного испытания обязательно произойдет одно и только одно событие данной группы.

Пример. Брошена игральная кость. Пусть $A_1, A_2, \ldots, A_6$ — события, состоящие в том, что на верхней грани брошенной кости выпало 1, 2, ..., 6 очков соответственно. События $A_1, A_2, \ldots, A_6$ образуют полную группу несовместных событий.

Среди всех возможных событий, которые в данном испытании могут произойти, можно выделить множество так называемых *элементарных событий* (или *элементарных исходов*), выделяемых из множества всех событий по следующим признакам:

1) все они взаимно исключают друг друга, и в результате испытания обязательно происходит одно из этих событий;

2) каково бы ни было событие A, по наступившему элементарному событию можно судить о том, что наступило или не наступило событие A.

Так, в примере с игральной костью множество элементарных событий состоит из шести элементарных событий $A_1, A_2, \ldots, A_6$, заключающихся в том, что на верхней грани брошенной кости выпало 1, 2, ..., 6 очков соответственно.

3.2. Операции над событиями. Два события A и B называют равными и пишут $A = B$, если наступление события A влечет за собой наступление события B и, наоборот, наступление события B влечет за собой наступление события A.

Пример. Брошены две игральные кости. Пусть событие A состоит в том, что сумма выпавших на костях очков четная, а событие B — в том, что на брошенных костях четности очков совпадают (т. е. либо на обеих костях выпало четное число очков, либо на обеих — нечетное). События A и B равны.

Объединением (или *суммой*) событий A и B называется событие, состоящее в наступлении события A или события B. Объединение событий обозначается

$$A \cup B \quad \text{или} \quad A + B.$$

Аналогично определяется сумма k событий $A_1, A_2, \ldots, A_k$, которая обозначается

$$\bigcup_k A_k.$$

Пересечением (или *произведением*) событий A и B называется событие, состоящее в одновременном наступлении как события A, так и события B. Пересечение событий A и B обозначается

$$A \cap B \quad \text{или} \quad A \cdot B.$$

Аналогично определяется произведение k событий $A_1, A_2, \ldots, A_k$, которое обозначается

$$\bigcap_k A_k.$$

Разностью событий A и B называется событие, состоящее в том, что наступает событие A, но не наступает событие B. Разность событий A и B обозначается

$$A \setminus B.$$

Сумма, произведение и разность событий геометрически могут быть изображены так же, как сумма, произведение и разность двух произвольных множеств.

3.3. Понятие вероятности. Рассмотрим опыт по бросанию монеты. Этот опыт имеет два взаимно исключающих друг друга исхода выпадение «герба» и выпадение «решетки». Обозначим эти исходы буквами A и B соответственно. Проведем серию испытаний, состоящую из n бросаний. Подсчитаем число выпадений «герба» и число выпадений «решетки». Повторим эту серию из n испытаний несколько раз, подсчитывая каждый раз число выпадений «герба» и «решетки». Пусть число испытаний, которые привели к выпадению «герба», равно $n(A)$. Отношение числа выпадения «герба» $n(A)$ к числу всех n испытаний данной серии опытов называют *частотой* события A в данной серии.

Как показывает практика, при больших n частоты $n(A)/n$ в различных сериях испытаний оказываются приблизительно одинаковыми и при достаточно больших n близкими к $1/2$.

Аналогичная ситуация будет и в случае бросания игральной кости: при достаточно большом числе испытаний в серии частота выпадения, скажем, шести очков в различных сериях испытаний будет примерно одинакова и близка к $1/6$.

Естественно, что как в примере с монетой, так и в примере с игральной костью полагалось, что монета и игральная кость геометрически правильны, сделаны из однородного материала и условия испытаний в различных сериях одинаковы.

Число подобных примеров можно продолжить, но во многих случаях при многократном повторении одного и того же опыта в одних и тех же условиях частота наступления некоторого события (т. е. отношение числа опытов, в которых этот результат наблюдался, к общему числу производимых испытаний) остается примерно одинаковой и близкой к некоторому числу P. Это число называют *вероятностью* рассматриваемого события.

Таким образом, вероятность того или иного события можно приближенно оценить по результатам серии опытов, состоящей из достаточно большого числа испытаний.

В так называемой классической модели теории вероятностей полагается, что:

1. Множество элементарных событий (исходов) одного испытания $\omega_1, \omega_2, \ldots, \omega_k$ конечно и образует полную группу несовместных событий.

2. С каждым элементарным исходом ω_i можно связать неотрицательное число p_i, называемое вероятностью этого исхода, причем $p_1+p_2+\ldots+p_k=1$.

3. Вероятность $P(A)$ наступления события A, состоящего в том, что наступило одно из событий ω_i, ω_j, ..., ω_s, равна сумме вероятностей наступления этих событий

$$P(A)=p_i+p_j+\ldots+p_s.$$

События ω_i, ω_j, ..., ω_s называют *благоприятствующими* событию A.

4. В случае, когда события ω_1, ω_2, ..., ω_k *равновероятны*, т. е. $p_1=p_2=\ldots=p_k=1/k$, вероятность $P(A)$ вычисляется по формуле

$$P(A)=r/k, \qquad (1)$$

где r — число исходов, благоприятствующих событию A, а k — число всех равновозможных исходов. Формула (1) выражает так называемое *классическое определение* вероятности.

Пример. Предположим, что брошены две игральные кости, грани которых занумерованы цифрами от 1 до 6. Какова вероятность выпадения «дубля»?

Каждый исход этого опыта можно представить как упорядоченную пару чисел $(m; n)$, где m — число очков, выпавших на первой кости, n — число очков на второй. Все исходы равновероятны, и число всевозможных различных исходов, т. е. число различных упорядоченных пар, равно 36. Событие A — выпадение «дубля» — происходит тогда и только тогда, когда наступает один из исходов $(m; n)$, при котором $m=n$. Таких исходов шесть. Следовательно, в согласии с формулой (1) вероятность события A

$$P(A)=6/36=1/6.$$

Сложение вероятностей. Рассмотрим несовместные события A_1 и A_2 и их объединение $A=A_1\cup A_2$. Пусть проводится серия испытаний, результатом каждого из которых могут быть события A, A_1 и A_2. Пусть n — число всех испытаний, $n(A)$, $n(A_1)$ и $n(A_2)$ — число тех испытаний, которые привели к наступлению событий A, A_1 и A_2 соответственно. Если в каком-либо испытании произошло событие A, то это значит, что произошло или событие A_1, или событие A_2 (одновременно события A_1 и A_2 произойти не могут, так как они несовместны). Поэтому числа $n(A)$, $n(A_1)$ и $n(A_2)$ связаны между собой равенством

$$n(A)=n(A_1)+n(A_2),$$

и, следовательно, частоты этих событий связаны равенством

$$\frac{n(A)}{n}=\frac{n(A_1)}{n}+\frac{n(A_2)}{n}.$$

При достаточно большом числе испытаний n частоты становятся близки к вероятностям соответствующих событий и, следовательно,

$$P(A)=P(A_1)+P(A_2).$$

Это равенство называется *формулой сложения вероятностей*.

Формула сложения вероятностей справедлива и в случае, когда событие A есть объединение любого конечного числа несовместных событий. Если $A=\bigcup\limits_k A_k$, то

$$P(A)=P(A_1)+P(A_2)+\ldots+P(A_k).$$

Из формулы сложения вероятностей, в частности, следует, что сумма вероятностей наступления события A и события, противоположного A (т. е. $\bar{A}$), равна единице:

$$P(A)+P(\bar{A})=1.$$

Вероятность произведения событий. Пусть A_1 и A_2 — независимые события, т. е. наступление события A_1 никак не отражается на условиях опыта, в результате которого может наступить событие A_2. Так, например, независимы события, состоящие в извлечении черного шара из двух различных урн, содержащих черные и белые шары. Однако два последовательных извлечения черного шара из одной урны (без возвращения вынутого шара обратно) не являются независимыми событиями (поскольку первое извлечение черного шара уменьшает число черных шаров, оставшихся в урне, и, следовательно, отражается на условиях второго опыта).

Пусть проводятся две серии испытаний: первая серия, состоящая из n испытаний, в результате которой $n(A_1)$ раз наступает событие A_1, и вторая серия, состоящая из m испытаний, в результате которой $m(A_2)$ раз наступает событие A_2. В результате этих двух серий испытаний событие $A=A_1\cap A_2$ наступит $n(A_1)\,m(A_2)$ раз, так как каждому из $n(A_1)$ исходов первого испытания соответствует $m(A_2)$ исходов второго испытания. Две проведенные серии испытаний всего могут иметь $n\cdot m$ исходов, так как каждому из n испытаний первой серии может соответствовать любое из m испытаний второй серии. Следовательно, частота наступления события $A=A_1\cap A_2$ равна

$$\frac{n(A_1)\,m(A_2)}{n\cdot m}=\frac{n(A_1)}{n}\cdot\frac{m(A_2)}{m}.$$

Так как при достаточно большом числе испытаний в каждой серии частоты становятся близки к вероятностям наступления событий, то вероятность наступления события $A=A_1\cap A_2$ равна

$$P(A)=P(A_1)\cdot P(A_2).$$

Для n независимых событий $A_1, A_2, \ldots, A_k$ вероятность их произведения вычисляется по формуле

$$P(A_1\cdot A_2\cdot\ldots\cdot A_k)=P(A_1)\cdot P(A_2)\cdot\ldots\cdot P(A_k).$$

Вероятность суммы совместных событий. Пусть A и B — два совместных события. Проведем серию из n испытаний. Пусть в этой серии испытаний событие A наступило $n(A)$ раз, а событие B — $n(B)$ раз. Допустим, что среди $n(A)+n(B)$ событий содержатся $n(A\cdot B)$ событий таких, при которых наступает как событие A, так и событие B. С другой стороны, так как событие $A+B$ наступает в том случае, когда наступает событие A, или событие B, или то и другое одновременно, то при n испытаниях событие $A+B$ наступит $n(A)+n(B)-n(A\cdot B)$ раз (в сумме $n(A)+n(B)$ имеется $n(A\cdot B)$ событий, которые засчитываются дважды). Следовательно, частота наступления события $A+B$ равна

$$\frac{n(A+B)}{n}=\frac{n(A)}{n}+\frac{n(B)}{n}-\frac{n(A\cdot B)}{n}.$$

Так как при достаточно большом числе испытаний частоты близки к вероятностям, то вероятность суммы двух совместных событий A

и B равна

$$P(A+B)=P(A)+P(B)-P(A\cdot B).$$

Пример. Найти вероятность того, что при одновременном бросании двух игральных костей хотя бы на одной кости выпадет 6 очков. Пусть A — событие, состоящее в том, что 6 очков выпало на первой игральной кости, B — событие, состоящее в том, что 6 очков выпало на второй кости. Так как события A и B совместны (6 очков может выпасть одновременно на первой и на второй костях), то вероятность указанного события равна

$$P(A+B)=P(A)+P(B)-P(A\cdot B)=\frac{1}{6}+\frac{1}{6}-\frac{1}{36}=\frac{11}{36}.$$

3.4. Условная вероятность. Рассмотрим следующий опыт. Пусть в урне находятся n шаров: m черных и $n-m$ белых. Пусть A — событие, состоящее в извлечении черного шара из урны, B — событие, состоящее в извлечении черного шара из той же урны после того, как из нее уже вынут один шар. Очевидно, что если событие A наступило, т. е. из урны первым был вынут черный шар, то в урне осталось $m-1$ черных шаров и вероятность наступления события B равна

$$\frac{m-1}{n-1}.$$

Если же первым был вынут белый шар (т. е. событие A не наступило), то в урне осталось m черных шаров и вероятность наступления события B равна

$$\frac{m}{n-1}.$$

Таким образом, вероятность наступления события B зависит от того, наступило или не наступило событие A.

Вероятность наступления события B при условии, что событие A произошло, называют *условной вероятностью события B при условии A* и обозначают

$$P(B\mid A).$$

Вероятность произведения зависимых событий вычисляется по формуле

$$P(A\cdot B)=P(A)\cdot P(B\mid A)=P(B)\cdot P(A\mid B), \qquad (2)$$

т. е. вероятность произведения двух зависимых событий A и B равна произведению вероятности одного из этих событий на условную вероятность другого, вычисленную при условии, что первое событие произошло.

Пример. Пусть в ящике находятся 5 изделий первого сорта и 3 изделия второго сорта. Из ящика последовательно вынимают два изделия. Найти вероятность того, что первое вынутое изделие первого сорта, а второе — второго сорта.

Пусть A — событие, состоящее в том, что первое вынутое изделие первого сорта, а B — событие, состоящее в том, что второе вынутое изделие — второго сорта.

Вероятность наступления события A

$$P(A)=5/8.$$

Так как в ящике осталось 4 изделия первого сорта и 3 изделия второго сорта, то вероятность вынуть изделие второго сорта равна

$$P(B\mid A)=3/7.$$

Искомое событие представляет собой произведение событий $A \cdot B$, и его вероятность вычисляется по формуле (2):

$$P(A \cdot B) = P(A) \cdot P(B \mid A) = \frac{5}{8} \quad \frac{3}{7} = \frac{15}{56}.$$

3.5. Формула полной вероятности. Для вычисления вероятности того или иного события A часто бывает удобно использовать *формулу полной вероятности*. Допустим, что имеется полная группа несовместных событий B_1, B_2, ..., B_k, и пусть A — некоторое событие, которое наступает при условии наступления некоторых из событий B_1, B_2, ..., B_k.

Так как события B_1, B_2, ..., B_k образуют полную группу несовместных событий, то событие A можно представить как сумму произведений событий:

$$A = A \cdot B_1 + A \cdot B_2 + \ldots + A \cdot B_k,$$

где все слагаемые в правой части равенства являются несовместными между собой событиями, и, следовательно, вероятность наступления события A равна сумме вероятностей событий $A \cdot B_1$, $A \cdot B_2$, ..., $A \cdot B_k$:

$$P(A) = P(A \cdot B_1) + P(A \cdot B_2) + \ldots + P(A \cdot B_k).$$

Применив к каждому из слагаемых, состоящих в правой части последнего равенства, формулу условной вероятности (2), получим формулу *полной вероятности*

$$P(A) = P(A \mid B_1) \cdot P(B_1) + P(A \mid B_2) \cdot P(B_2) + \ldots + P(A \mid B_k) \cdot P(B_k).$$

Пример. По цели произведено три последовательных выстрела. Вероятность попадания при первом выстреле равна $P_1 = 0{,}6$, при втором — $P_2 = 0{,}7$, при третьем — $P_3 = 0{,}9$. Известно, что при одном (любом из трех) попадании вероятность поражения цели равна 0,4, при двух (любых из трех) — 0,8, при трех — 1,0 (т. е. цель обязательно будет поражена). Найти вероятность поражения цели при трех выстрелах.

Пусть A — событие, состоящее в том, что цель поражена; B_1, B_2, B_3 — события, состоящие в том, что были одно, два и три попадания в цель соответственно; B_4 — событие, состоящее в том, что в цель не попали ни разу. Очевидно, что события B_1, B_2, B_3, B_4 образуют полную группу несовместных событий (одно из них обязательно произойдет, и наступление одного события исключает наступление другого).

Вероятность наступления события A вычислим по формуле полной вероятности:

$$P(A) = P(A \mid B_1) \cdot P(B_1) + P(A \mid B_2) \cdot P(B_2) + P(A \mid B_3) \cdot P(B_3) + P(A \mid B_4) \cdot P(B_4).$$

По условию задачи вероятность поражения цели при одном попадании $P(A \mid B_1) = 0{,}4$, при двух — $P(A \mid B_2) = 0{,}8$, при трех — $P(A \mid B_3) = 1{,}0$.

Вычислим вероятности наступления событий B_1, B_2 и B_3.

Событие B_1 может наступить в следующих трех случаях.

1. Первый выстрел попал в цель, а второй и третий — не попали. Вероятность этого события можно вычислить как вероятность произведения указанных трех независимых событий:

$$P_1 (1 - P_2)(1 - P_3) = 0{,}6 \cdot 0{,}3 \cdot 0{,}1 = 0{,}018,$$

где $1 - P_2$ и $1 - P_3$ — вероятности промахов при втором и третьем выстрелах.

2. Второй выстрел попал в цель, а первый и третий — нет. Вероятность этого события вычисляется аналогично предыдущей и равна

$$(1 - P_1)\, P_2\, (1 - P_3) = 0{,}4 \quad 0{,}7 \cdot 0{,}1 = 0{,}028.$$

3. Только третий выстрел попал в цель. Вероятность этого события равна

$$(1-P_1)\cdot(1-P_2)\cdot P_3=0{,}4\cdot 0{,}3\cdot 0{,}9=0{,}108.$$

Таким образом, вероятность наступления события B_1 равна

$$P(B_1)=0{,}018+0{,}028+0{,}108=0{,}154.$$

Аналогично вероятность наступления события B_2 (двух попаданий при трех выстрелах) равна

$$P(B_2)=P_1\,P_2\cdot(1-P_3)+P_1\,(1-P_2)\cdot P_3+(1-P_1)\cdot P_2\cdot P_3=$$
$$=0{,}6\;0{,}7\cdot 0{,}1+0{,}6\cdot 0{,}3\cdot 0{,}9+0{,}4\cdot 0{,}7\cdot 0{,}9=0{,}042+0{,}162+0{,}252=0{,}456.$$

И, наконец, вероятность наступления события B_3 равна

$$P(B_3)=P_1\,P_2\cdot P_3=0{,}6\cdot 0{,}7\cdot 0{,}9=0{,}378.$$

Вероятность наступления четвертого события подсчитывать не будем, так как очевидно, что вероятность поражения цели при трех промахах равна нулю:

$$P(A\mid B_4)=0.$$

Подставив полученные значения вероятностей в формулу полной вероятности, находим вероятность поражения цели

$$P(A)=0{,}4\cdot 0{,}154+0{,}8\cdot 0{,}456+1{,}0\cdot 0{,}378=0{,}8044.$$

3.6. Формула Бернулли. Пусть производится n независимых испытаний, в каждом из которых вероятность наступления события A равна P (а вероятность события $\overline{A}$, противоположного событию A, равна $1-P$). Требуется найти вероятность того, что при n испытаниях событие A произойдет m раз.

Событие B, состоящее в том, что событие A наступило при каждом из m первых испытаний и не произошло при остальных $n-m$ испытаниях, можно записать в виде произведения m событий A и $n-m$ событий $\overline{A}$:

$$B=\underbrace{A\cdot A\cdot\ldots\cdot A}_{m}\cdot\underbrace{\overline{A}\cdot\overline{A}\cdot\ldots\cdot\overline{A}}_{n-m}. \tag{3}$$

Так как все n испытаний по условию независимы, то вероятность наступления события B равна произведению вероятностей наступления событий A и $\overline{A}$:

$$P(B)=P^m\cdot(1-P)^{n-m}.$$

Событие A может произойти m раз при n испытаниях, но при этом может получиться последовательность чередований событий, отличная от последовательности (3). Однако для любой такой последовательности событий вероятность наступления события A m раз при n испытаниях равна $P^m\cdot(1-P)^{n-m}$. Число всевозможных различных последовательностей из n элементов, содержащих m элементов A, равно числу сочетаний C_n^m из n элементов по m элементов. Так как различные последовательности из m элементов A и $n-m$ элементов $\overline{A}$ могут быть рассмотрены как различные исходы серии из n независимых испытаний, то все эти исходы (события) являются несовместными и искомая вероятность, обычно обозначаемая $P_{m,n}$, вычисляется как сумма вероятностей наступления различных исходов серии из n испытаний:

$$P_{m,n}=C_n^m P^m(1-P)^{n-m}.$$

Эта формула называется *формулой Бернулли*.

Пример. В урне находятся 15 шаров: 10 черных и 5 белых. Испытание состоит в следующем: вынимают один шар, фиксируют его цвет и кладут обратно; перемешивают шары и вынимают еще один шар. Указанную процедуру повторяют некоторое число раз. Какова вероятность того, что при таком испытании из пяти вынутых шаров будет четыре белых?

Вероятность появления белого шара при каждом испытании равна $P=5/15=1/3$, а вероятность того, что белый шар не будет вытащен, равна $1-P=2/3$. По формуле Бернулли находим вероятность появления четырех белых шаров среди пяти вынутых:

$$P_{4,5}=C_5^4\cdot P^4\cdot(1-P)^{5-4}=C_5^4P^4(1-P)=5\cdot\left(\frac{1}{3}\right)^4\cdot\left(\frac{2}{3}\right)=\frac{10}{243}.$$

ПРИЛОЖЕНИЯ

I. ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ

К элементарным функциям относятся рациональные, степенные, показательная и логарифмическая функции, а также тригонометрические функции. К классу элементарных функций, кроме того, относят также сложные функции, образованные из перечисленных выше элементарных функций.

Показательная, логарифмическая, тригонометрические функции, а также сложные функции, образованные с помощью перечисленных функций, называют элементарными трансцендентными функциями.

§ 1. График функции и простейшие преобразования графика функции

Графиком функции $y=f(x)$ в прямоугольной декартовой системе координат Oxy называют множество всех упорядоченных пар точек с координатами $(x;\ f(x))$.

Простейшими преобразованиями графика функции $y=f(x)$ в прямоугольной декартовой системе координат Oxy являются перечисленные ниже преобразования.

1. Перенос графика параллельно оси ординат. График функции $y=f(x)+a$ получается из графика функции $y=f(x)$ параллельным переносом $r=(0;\ a)$. Если число a положительно, то график функции $y=f(x)+a$ получается сдвигом графика функции $y=f(x)$ на a единиц вверх; если a отрицательно — на a единиц вниз (рис. П.1).

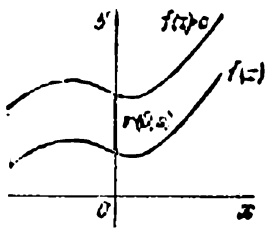


Рис. П.1

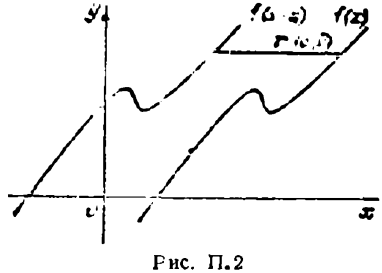


Рис. П.2

2. Перенос графика параллельно оси абсцисс. График функции $y=f(x-a)$ получается из графика функции $y=f(x)$ параллельным переносом $r=(a;\ 0)$. Если $a<0$, то график переносится влево, если $a>0$ — вправо (рис. П.2).

3. Растяжения и сжатия графиков функции. Пусть дан график функции $y=f(x)$ и точка с координатами $(x_0;\ y_0)$ принадлежит графику функции $y=f(x)$.

Рассмотрим функцию $y = kf(x)$ (k — положительное число). Точка графика функции $y = kf(x)$ с абсциссой x_0 имеет ординату ky_0. При $k > 1$ ординаты точек графика функции $y = kf(x)$ увеличиваются в k раз по сравнению с ординатами точек графика функции $y = f(x)$, имеющих ту же самую абсциссу. В этом случае говорят, что график

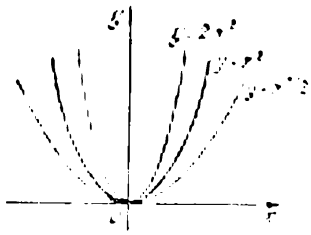


Рис. П.3

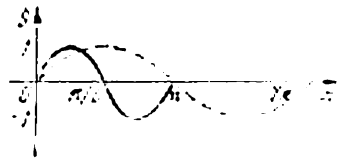


Рис. П.4

функции $y = kf(x)$ получается из графика функции $y = f(x)$ растяжением от оси абсцисс в k раз.

При $0 < k < 1$ ординаты точек графика функции $y = kf(x)$ уменьшаются в $1/k$ раз по сравнению с ординатами точек графика функции $y = f(x)$, имеющих ту же самую абсциссу. В этом случае говорят, что график функции $y = kf(x)$ получается из графика функции $y = f(x)$ сжатием к оси абсцисс в k раз.

Например, график функции $y = 2x^2$ получается из графика функции $y = f(x)$ растяжением от оси абсцисс в 2 раза, а график функции $y = \frac{1}{2}x^2$ — сжатием к оси абсцисс в 2 раза (рис. П.3).

Рассмотрим функцию $y = f(kx)$ (k — положительное число). Точка графика функции $y = f(kx)$ с ординатой y_0 имеет абсциссу x_0/k. При $k > 1$ абсциссы всех точек графика функции $y = f(kx)$ уменьшаются в k раз по сравнению с абсциссами точек графика функции $y = f(x)$, имеющих ту же ординату. В этом случае говорят, что график функции $y = f(kx)$ получается из графика функции $y = f(x)$ сжатием к оси ординат в k раз.

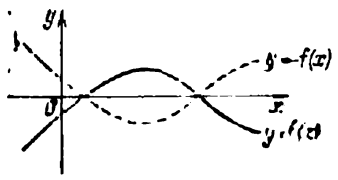


Рис. П.5

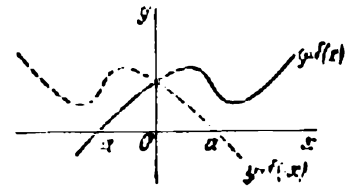


Рис. П.6

При $0 < k < 1$ абсциссы точек графика функции $y = f(kx)$ увеличиваются в $1/k$ раз по сравнению с абсциссами точек графика функции $y = f(x)$, имеющих ту же самую ординату. В этом случае говорят, что график функции $y = f(kx)$ получается из графика функции $y = f(x)$ растяжением от оси ординат в $1/k$ раз. Например, график функции $y = \sin 2x$ получается из графика функции $y = \sin x$ сжатием к оси ординат в 2 раза (рис. П.4).

4. График функции $y = -f(x)$ получается из графика функции $y = f(x)$ симметричным отражением относительно оси Ox (рис. П.5).

5. График функции $y=f(-x)$ получается из графика функции $y=f(x)$ симметричным отражением относительно оси Oy (рис. П.6).

Примеры. 1. С помощью перечисленных выше преобразований графиков функций получить из графика функции $y=x^2$ график функции $y=2x^2+4x-1$.

Запишем квадратный трехчлен $2x^2+4x-1$ в виде $2(x+1)^2-3$, график которого (рис. П.7) получается из графика квадратного трехчлена $y=x^2$ в результате следующих преобразований:

а) параллельного переноса $\mathbf{r}=(-1;\ 0)$;

б) растяжения от оси абсцисс в 2 раза;

в) параллельного переноса $\mathbf{r}=(0;\ -3)$.

2. С помощью преобразований графиков функций получить из графика функции $y=\sin x$ график функции $y=2\sin\left(x+\frac{\pi}{4}\right)+\frac{1}{2}$.

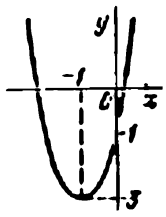


Рис. П.7

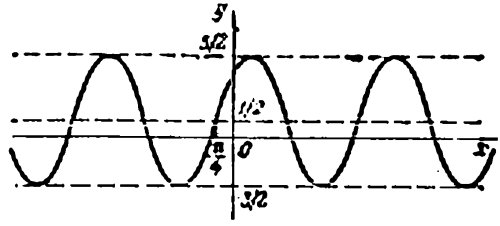


Рис. П.8

График этой функции (рис. П.8) получается из графика функции $y=\sin x$ в результате следующих преобразований:

а) параллельного переноса $\mathbf{r}=(-\pi/4;\ 0)$;

б) растяжения от оси абсцисс в 2 раза;

в) параллельного переноса $\mathbf{r}=\left(0;\ \frac{1}{2}\right)$.

§ 2. Линейная функция

Линейной функцией называется функция вида

$$f(x)=kx+b,$$

где k и b — некоторые числа.

Свойства линейной функции.

1. Область определения — множество всех действительных чисел $\mathbf{R}$.

2. Область изменения (множество значений) при $k\neq 0$ — множество всех действительных чисел. При $k=0$ множество значений функции состоит из одной точки b.

3. При $k\neq 0$ и $b\neq 0$ функция не является ни четной, ни нечетной. При $k=0$ (b — любое) функция четная, при $b=0$ и $k\neq 0$ — нечетная.

4. Линейная функция непрерывна и дифференцируема на всей числовой оси; ее производная в каждой точке равна k.

5. Линейная функция не имеет экстремумов ни при каких значениях k и b. При $k\neq 0$ критических точек нет. При $k=0$ каждая точка является критической точкой функции.

6. При $k > 0$ линейная функция возрастает при всех $x \in \mathbb{R}$ при $k < 0$ — убывает при всех $x \in \mathbb{R}$, при $k = 0$ постоянна.

7. График линейной функции пересекает ось Oy в точке $y = b$. При $k \neq 0$ график пересекает ось Ox в точке $x = -b/k$, при $k = 0$ он параллелен оси Ox.

На рис. П.9, П.10 изображены графики линейных функций.

Линейная функция $f(x) = kx + b$ может иметь своим графиком любую прямую координатной плоскости Oxy за исключением вертикальных прямых.

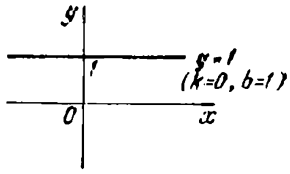


Рис. П.9

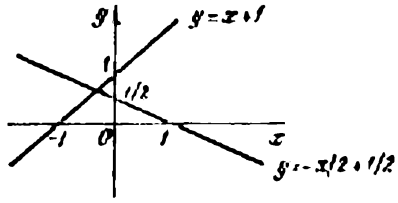


Рис. П.10

Прямая пропорциональность. Переменную y называют *прямо пропорциональной* переменной x с коэффициентом пропорциональности k, если эти переменные связаны соотношением $y = kx$, где k — некоторое действительное число, отличное от нуля.

Если считать x независимой переменной, а y — зависимой, то уравнение $y = kx$ определяет y как функцию x. Графиком этой функции является прямая, проходящая через начало координат, с угловым коэффициентом k (тангенс угла наклона прямой к оси Ox равен k). Прямо пропорциональная зависимость является частным случаем линейной функции.

§ 3. Обратно пропорциональная зависимость

Переменную y называют *обратно пропорциональной* переменной x, если значения этих переменных связаны равенством $y = k/x$, где k — некоторое действительное число, отличное от нуля. Число k называют *коэффициентом обратной пропорциональности*.

Если считать x независимой переменной, а y — зависимой, то уравнение $y = k/x$ определяет y как функцию от x. График функции $y = k/x$ называется *гиперболой*.

Свойства функции $f(x) = k/x$.

1. Область определения — множество всех действительных чисел, за исключением числа 0.

2. Область изменения (множество значений) — множество всех действительных чисел, за исключением числа 0.

3. Функция $f(x) = k/x$ нечетная, и ее график симметричен относительно начала координат. Она непрерывна и дифференцируема во всей области определения: $f'(x) = -k/x^2$, критических точек не имеет.

4. Функция $f(x) = k/x$ при $k > 0$ монотонно убывает на промежутках в $(-\infty, 0)$ и $(0, +\infty)$, а при $k < 0$ монотонно возрастает на тех же промежутках.

5. График функции $y = k/x$ при $k > 0$ в промежутке $(0; +\infty)$ направлен вогнутостью вверх, а в промежутке $(-\infty, 0)$ — вогнуто-

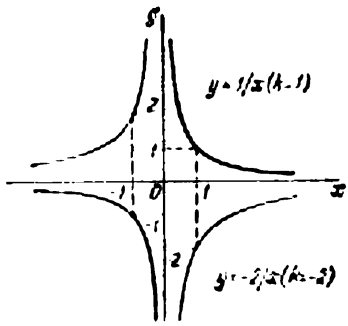


Рис. П.11

стью вниз. При $k < 0$ промежуток вогнутости вверх — $(-\infty; 0)$, а промежуток вогнутости вниз — $(0; +\infty)$.

Графики функции $y = k/x$ при $k = 1$ и $k = -2$ изображены на рис. П.11.

§ 4. Дробно-линейная функция

Функция вида $f(x) = \dfrac{ax+b}{cx+d}$ называется *дробно-линейной* функцией (a, b, c, d — некоторые числа, $c \neq 0$) *). Преобразуем дробно-линейное выражение

$$\frac{ax+b}{cx+d} = \frac{a}{c} \cdot \frac{x+\dfrac{b}{a}}{x+\dfrac{d}{c}} = \frac{a}{c} \cdot \left[1 + \frac{\dfrac{b}{a} - \dfrac{d}{c}}{x+\dfrac{d}{c}}\right] = \frac{\dfrac{b}{c} - \dfrac{da}{c^2}}{x+\dfrac{d}{c}} + \frac{a}{c}.$$

График дробно-линейной функции $y = \dfrac{ax+b}{cx+d}$ может быть получен из графика обратно пропорциональной зависимости следующими преобразованиями:

а) параллельным переносом $\mathbf{r} = (-d/c;\ 0)$;

б) сжатием (или растяжением) к оси абсцисс в $\left(\dfrac{b}{c} - \dfrac{da}{c^2}\right)$ раз;

в) параллельным переносом $\mathbf{r} = (0;\ a/c)$.

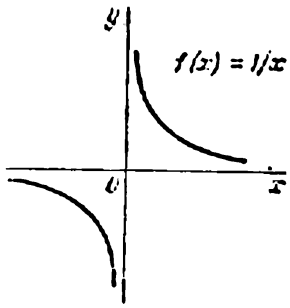


Рис. П.12

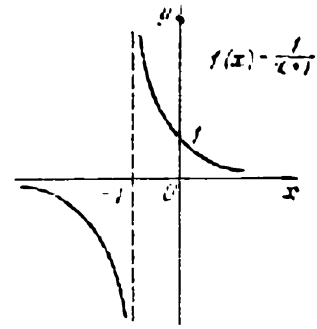


Рис. П.13

Последовательное преобразование графика функции $y = \dfrac{1}{x}$ в график функции $y = \dfrac{2x+1}{x+1}$ приведено на рис. П.12–П.15. На рис. П.12 изображен график функции $y = \dfrac{1}{x}$. На рис. П.13 изображен график функции $y = \dfrac{1}{x+1}$, получающейся из графика функции $y = \dfrac{1}{x}$ параллельным переносом $\mathbf{r} = (-1;\ 0)$. На рис. П.14

*) Если $c = 0$, то дробно-линейная функция превращается в линейную.

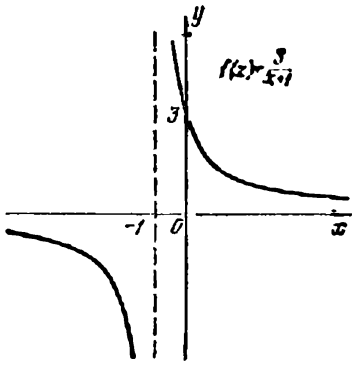


Рис. П.14

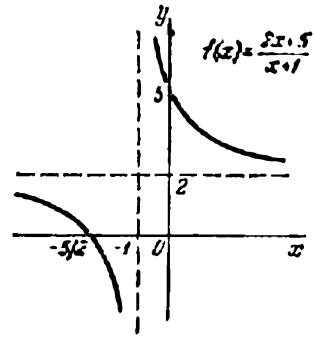


Рис. П.15

изображен график функции $y=\frac{3}{x+1}$, полученный из графика $y=\frac{1}{x+1}$ растяжением от оси абсцисс в 3 раза. И, наконец, на рис. П.15 изображен график дробно-линейной функции $y=\frac{2x+5}{x+1}$, который получается из графика $y=\frac{3}{x+1}$ параллельным переносом $r=(0;\ 2)$.

§ 5. Квадратичная функция

Функция $f(x)=ax^2+bx+c$, где a, b, c — некоторые действительные числа ($a\neq 0$), называется *квадратичной функцией*. График квадратичной функции называется *параболой*.

Квадратичная функция может быть приведена к виду

$$f(x)=a\left(x+\frac{b}{2a}\right)^2-\frac{b^2-4ac}{4a}. \qquad (1)$$

Представление квадратичной функции в виде (1) называется *выделением полного квадрата*.

Свойства квадратичной функции.

1. Область определения — вся числовая прямая.

2. При $b\neq 0$ функция не является четной и не является нечетной. При $b=0$ квадратичная функция четная.

3. Квадратичная функция непрерывна и дифференцируема во всей области определения.

4. Функция имеет единственную критическую точку $x=-b/(2a)$. Если $a>0$, то в точке $x=-b/(2a)$ функция имеет минимум. При $x<-b/(2a)$ функция монотонно убывает, при $x>-b/(2a)$ — монотонно возрастает.

Если $a<0$, то в точке $x=-b/(2a)$ функция имеет максимум. При $x<-b/(2a)$ функция монотонно возрастает, при $x>-b/(2a)$ — монотонно убывает.

Точка графика квадратичной функции с абсциссой $x=-b/(2a)$ и ординатой $y=-\frac{b^2-4ac}{4a}$ называется *вершиной параболы*.

5. Область изменения функции: при $a>0$ множество значений функции — промежуток $\left[-\frac{b^2-4ac}{4a};\ +\infty\right)$; при $a<0$ множество значений функции — промежуток $\left(-\infty;\ -\frac{b^2-4ac}{4a}\right]$.

6. График квадратичной функции пересекается с осью Oy в точке $y=c$. В случае, если $b^2-4ac>0$, график квадратичной функции пересекает ось Ox в двух точках (которые являются различными действительными корнями квадратного уравнения); если $b^2-4ac=0$ (квадратное уравнение имеет один корень кратности 2), график квадратичной функции касается оси Ox в точке $x=-b/(2a)$; если $b^2-4ac<0$, пересечения с осью Ox нет.

Из представления квадратичной функции в виде (1) также следует, что график функции симметричен относительно прямой $x=-b/(2a)$.

График функции $y=ax^2+bx+c$, записанной в виде (1), может быть получен из графика функции $y=x^2$ следующими преобразованиями:

а) параллельным переносом $\mathbf{r}=(-b/(2a);\ 0)$;

б) сжатием (или растяжением) к оси абсцисс в a раз;

в) параллельным переносом $\mathbf{r}=\left(0;\ -\frac{b^2-4ac}{4a}\right)$.

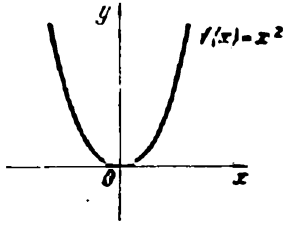


Рис. П.16

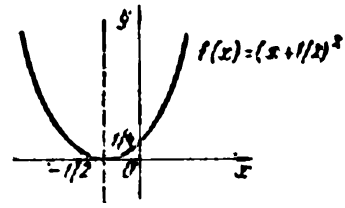


Рис. П.17

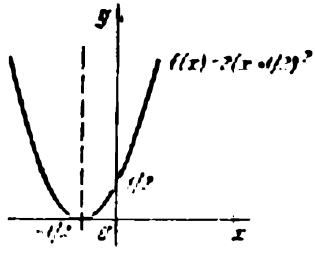


Рис. П.18

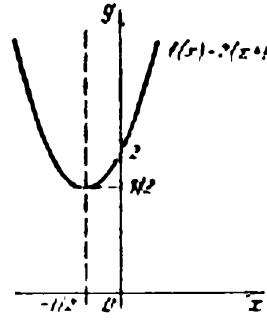


Рис. П.19

Пример. Построить график функции

$$y=2x^2+2x+2.$$

Приведем квадратный трехчлен к виду (1):

$$f(x)=2x^2+2x+2=2\left(x+\frac{1}{2}\right)^2+\frac{3}{2}.$$

График функции $y=2x^2+x+2$ может быть получен из графика функции $y=x^2$ следующими преобразованиями:

а) параллельным переносом $\mathbf{r}=(-1/2;\ 0)$;

б) растяжением от оси абсцисс в 2 раза;

в) параллельным переносом $\mathbf{r}=(0;\ 3/2)$.

Результаты последовательного применения указанных преобразований изображены на рис. П.16–П.19.

§ 6. Степенная функция

Степенной функцией называется функция вида $f(x) = x^\alpha$, где α — любое действительное число, называемое *показателем степени*.

Свойства степенной функции.

1. Область определения — множество всех положительных чисел.
2. Область изменения — множество всех положительных чисел.

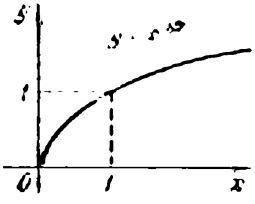


Рис. II.20

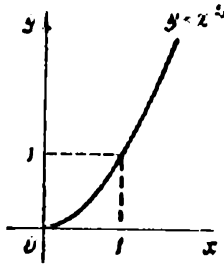


Рис. II.21

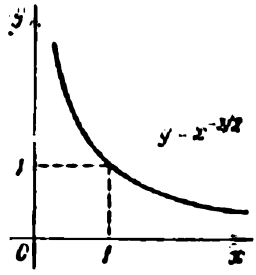


Рис. II.22

3. Степенная функция не является четной и не является нечетной.
4. Степенная функция непрерывна во всей области определения.
5. Степенная функция дифференцируема во всей области определения и ее производная вычисляется по формуле

$$(x^\alpha)' = \alpha \cdot x^{\alpha-1}.$$

6. Степенная функция x^α монотонно возрастает во всей области определения при $\alpha > 0$ и монотонно убывает при $\alpha < 0$.
7. При $\alpha < 0$ и $\alpha > 1$ график степенной функции направлен вогнутостью вверх, а при $0 < \alpha < 1$ — вогнутостью вниз.

Графики степенной функции при $\alpha = 1/2$, $\alpha = 5/2$ и $\alpha = -3/2$ приведены на рис. II.20—II.22 *).

§ 7. Показательная функция

Показательной функцией называется функция вида $f(x) = a^x$, где a — некоторое положительное действительное число, называемое *основанием степени*. При $a = 1$ значение показательной функции при любом значении аргумента равно единице, и случай $a = 1$ далее не будет рассматриваться.

Свойства показательной функции.

1. Область определения — вся числовая прямая.
2. Область изменения (множество значений) — множество всех положительных чисел.
3. Функция непрерывна и дифференцируема во всей области определения, ее производная вычисляется по формуле

$$(a^x)' = a^x \ln a.$$

*) Здесь степенная функция определяется только на положительной полуоси. Степенную функцию при некоторых конкретных α можно определить на всей числовой прямой.

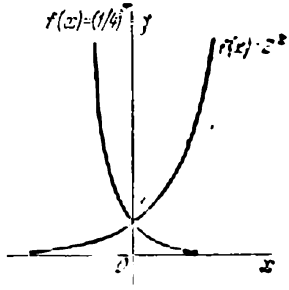


Рис. П.23

4. При $a > 1$ функция монотонно возрастает, при $a < 1$ — монотонно убывает.

5. Показательная функция имеет обратную функцию, называемую логарифмической функцией.

6. График любой показательной функции пересекает ось Oy в точке $y = 1$.

7. График показательной функции — кривая, направленная вогнутостью вверх.

Графики показательной функции при значениях $a = 2$ и $a = 1/4$ изображены на рис. П.23.

§ 8. Логарифмическая функция

Функцию, обратную показательной функции $y = a^x$, называют *логарифмической* и обозначают

$$y = \log_a x.$$

Число a называется *основанием* логарифмической функции. Логарифмическую функцию с основанием 10 обозначают $\lg x$, а логарифмическую функцию с основанием e обозначают $\ln x$.

Свойства логарифмической функции.

1. Область определения — промежуток $(0; +\infty)$.

2. Область изменения (множество значений) — вся числовая прямая.

3. Логарифмическая функция непрерывна и дифференцируема во всей области определения, ее производная вычисляется по формуле

$$(\log_a x)' = \frac{1}{x \ln a}.$$

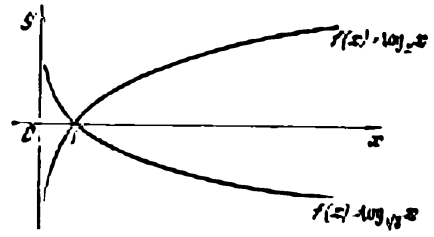


Рис. П.24

4. Логарифмическая функция монотонно возрастает, если $a > 1$, и монотонно убывает, если $0 < a < 1$.

5. Графики всех логарифмических функций проходят через точку с координатами (1; 0).

6. При $a > 1$ график логарифмической функции — кривая, направленная вогнутостью вниз; при $0 < a < 1$ — кривая, направленная вогнутостью вверх.

Графики логарифмической функции при $a = 2$ и $a = 1/3$ изображены на рис. П.24.

II. СИСТЕМЫ СЧИСЛЕНИЯ

Наряду с широко используемой десятичной позиционной системой счисления, при работе на электронно-вычислительных машинах обычно используются двоичные, восьмеричные, шестнадцатеричные (а также смешанные) системы счисления. Излагая основные сведения о восьмеричных и двоичных системах счисления и способах перевода чисел из одной системы счисления в другую, вместо слов «число,

записанное в десятичной (восьмеричной, двоичной) системе счисления», будем просто говорить «десятичное (соответственно восьмеричное, двоичное) число».

§ 1. Восьмеричная система счисления

В восьмеричной позиционной системе счисления для записи чисел используются восемь различных цифр: 0, 1, 2, 3, 4, 5, 6, 7. Число восемь записывается как 10. В восьмеричной позиционной системе счисления запись

$$a_n a_{n-1} \ldots a_1 a_0,\ a_{-1} \ldots a_{-k} \ldots, \tag{1}$$

где $a_n, a_{n-1}, \ldots, a_1, a_0, a_{-1}, \ldots, a_{-k}, \ldots$ — цифры восьмеричной системы счисления и хотя бы одна из цифр отлична от нуля, представляет собой сокращенную запись положительного числа a:

$$a = a_n \cdot 8^n + a_{n-1} \cdot 8^{n-1} + \ldots + a_1 \cdot 8^1 + a_0 \cdot 8^0 + a_{-1} \cdot 8^{-1} + \ldots$$
$$\ldots + a_{-k} \cdot 8^{-k} + \ldots \tag{2}$$

Для записи отрицательного числа $-a$ перед последовательностью цифр (1) ставят знак минус. Все понятия, введенные для десятичных дробей, являющихся формой записи чисел в десятичной позиционной системе счисления, дословно переносятся и на выражения вида (1), которые можно назвать восьмеричными дробями.

Правила перевода десятичных дробей в восьмеричные.

Перевод целой части числа. Разделим целую часть десятичного числа на число 8. Цифра, полученная в остатке, есть последняя цифра восьмеричной записи данного числа. Полученное частное делим на число 8; цифра, являющаяся остатком, — предпоследняя цифра восьмеричной записи числа. Продолжая процесс деления получаемых частных, последовательно выписываем цифры восьмеричной записи числа. Процесс деления прекращается, когда в частном получится цифра, меньшая 8, которая будет первой цифрой восьмеричной записи данного числа.

Пример 1. Записать десятичное число 201 в восьмеричной системе счисления:

$$\begin{array}{r|l} 201 & 8 \\ \underline{-16} & \overline{25 \mid 8} \\ 41 & \underline{-24 \mid 3} \\ \underline{40} & 1 \\ 1 & \end{array}$$

Процесс деления частных закончен на числе 3, так как $3 < 8$. Десятичное число 201 в восьмеричной системе счисления записывается как 311.

Перевод дробной части числа. Умножаем дробную часть числа на 8. Целая часть произведения является первой цифрой после запятой в восьмеричной записи дробной части числа. Умножая дробную часть полученного произведения на 8, получаем число, целая часть которого — вторая цифра восьмеричной записи дроби. Процесс умножения продолжается до тех пор, пока дробная часть произведения не будет состоять из одних нулей, либо до тех пор, пока не получим требуемого количества цифр восьмеричной записи дробной части числа. При этом может оказаться, что конечная деся-

тичная дробь записывается в виде бесконечной периодической восьмеричной дроби (см. пример 3).

Пример 2. Записать десятичное число 0,1875 в восьмеричной системе счисления.

Процедуру умножения изобразим в виде двух столбцов чисел, где левый столбец – целая часть числа, которая при выполнении следующего умножения заменяется на нуль:

0	1875
×	8
1	5000
×	8
4	0000

Десятичное число 0,1875 записывается в виде восьмеричной дроби 0,14.

Пример 3. Записать десятичную дробь 0,3 в восьмеричной системе счисления.

Перевод десятичной дроби 0,3 в восьмеричную систему счисления будет осуществляться так же, как и в предыдущем примере:

0	3
×	8
2	4
×	8
*3	2
×	8
1	6
×	8
4	8
×	8
6	4
×	8
*3	2
×	8

Нетрудно заметить, что начиная с шага умножения, отмеченного второй звездочкой, в левом столбце будет появляться комбинация цифр 3146, представляющая собой период восьмеричной дроби. Таким образом, десятичная дробь 0,3 в восьмеричной системе счисления записывается в виде бесконечной периодической дроби 0,2 (3146).

Перевод чисел, записанных в восьмеричной системе счисления, в десятичную систему обычно производят с помощью формулы (2).

Пример 4. Записать восьмеричное число 163,2 в десятичной системе счисления.

В согласии с формулой (2) запись 163,2 есть сокращенная запись числа

$$1\cdot 8^2+6\cdot 8^1+3\cdot 8^0+2\cdot 8^{-1}=64+48+3+1/4=64+48+3+0{,}25=115{,}25.$$

Таким образом, восьмеричное число 163,2 в десятичной системе счисления записывается как 115,25.

§ 2. Двоичная система счисления

В двоичной позиционной системе счисления для записи чисел используются два символа (цифры): 0 (нуль) и 1 (единица). Число два в двоичной системе счисления записывается как 10. В двоичной системе счисления запись

$$a_n a_{n-1} \ldots a_1 a_0, \; a_{-1} \ldots a_{-k} \ldots \tag{3}$$

где a_n, a_{n-1}, ..., a_1, a_0, a_{-1}, ..., a_{-k}, ... — цифры двоичной системы счисления (0 или 1) и хотя бы одна из цифр отлична от нуля, представляет собой сокращенную запись положительного числа a:

$$a = a_n \cdot 2^n + a_{n-1} \cdot 2^{n-1} + \ldots + a_1 \cdot 2^1 + a_0 \cdot 2^0 + a_{-1} \cdot 2^{-1} + \ldots$$
$$\ldots + a_{-k} \cdot 2^{-k} + \ldots \tag{4}$$

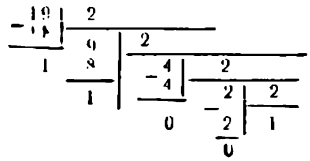
Для записи отрицательного числа $-a$ перед последовательностью цифр (3) ставится знак минус. Все понятия, вводимые для десятичных дробей, дословно переносятся на выражения вида (3).

Двоичные таблицы сложения и умножения имеют следующий вид.

Таблица сложения	Таблица умножения
$0+0=0$	$0\times0=0$
$0+1=1$	$0\times1=0$
$1+0=1$	$1\times0=0$
$1+1=10$	$1\times1=1$

Правила перевода чисел десятичной системы счисления в двоичную аналогичны правилам перевода из десятичной системы счисления в восьмеричную.

Пример 5. Записать десятичное число 19 в двоичной системе счисления



Таким образом, число 19 в двоичной системе счисления записывается как 10 011.

Перевод чисел из десятичной системы счисления в двоичную можно произвести и другим способом. Этот способ основан на том, что между числами 0, 1, 2, 3, 4, 5, 6, 7 и двоичными числами, состоящими из трех цифр, существует взаимно однозначное соответствие, а именно перечисленные числа в двоичной системе счисления записываются соответственно как

000, 001, 010, 011, 100, 101, 110, 111.

Основываясь на существовании такого соответствия, десятичное число удобнее переводить в двоичную систему в два этапа: сначала записать десятичное число в восьмеричной системе счисления, потом каждую восьмеричную цифру заменить соответствующей группой из трех двоичных цифр.

Пример 6. Записать десятичное число 201,1875 в двоичной системе счисления.

Данное десятичное число в восьмеричной системе счисления запишется как 311,14 (см. примеры 1, 2). Каждую цифру восьмеричного числа 311,14 заменяем группой из трех двоичных цифр:

$$3 \to 011; \quad 1 \to 001; \quad 4 \to 100.$$

В результате восьмеричное число 311,14 в двоичной системе счисления записывается в виде

$$011001001,001100.$$

Отбрасывая нули, стоящие в начале целой и конце дробной части двоичного числа, окончательно получаем, что десятичное число 201, 1875 записывается в двоичной системе счисления в виде 11001001,0011.

Перевод чисел из двоичной системы счисления в десятичную удобно производить с помощью формулы (4).

Пример 7. Записать двоичное число 10110,11 в десятичной системе счисления.

В согласии с формулой (4) запись 10110,11 есть сокращенная запись числа

$$1\cdot 2^4+0\cdot 2^3+1\cdot 2^2+1\cdot 2^1+0\cdot 2^0+1\cdot 2^{-1}+1\cdot 2^{-2}=$$

$$=16+4+2+\frac{1}{2}+\frac{1}{4}=22+\frac{3}{4}=22{,}75.$$

Итак, двоичное число 10110,11 в десятичной системе счисления записывается как 22,75.

СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ

$[a, b]$ — закрытый промежуток (отрезок): точки (числа) a и b принадлежат промежутку

$(a; b]$ и $[a; b)$ — полуоткрытые промежутки: точка a не принадлежит промежутку; точка принадлежит промежутку; соответственно, точка a принадлежит, а b не принадлежит промежутку

$(a; b)$ — открытый промежуток (интервал): точки a и b не принадлежат промежутку

$\Rightarrow$ — символ, означающий «следует в одну сторону»

$\Leftrightarrow$ — символ, означающий «следует в обе стороны»

$\in$ — символ принадлежности

$\notin$ — символ «не принадлежит»

$\varnothing$ — символ пустого множества

$\cup$ — символ объединения множеств

$\cap$ — символ пересечения множеств

$\setminus$ — символ разности и дополнения множеств

$\sim$ — знак эквивалентности 18 (в геометрии — символ подобия)

f и f^{-1}, g и g^{-1} и т. д. — обозначения взаимно обратных отображений

$n!$ — обозначение произведения $1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$ (читается «эн факториал»)

P_n — число перестановок из n элементов

C_n^m или $\binom{n}{m}$ — число сочетаний из n элементов по m элементов

A_n^m — число размещений из n элементов по m элементов

$\|a_{ij}\|$ или (a_{ij}) или $[a_{ij}]$ — матрица, составленная из чисел a_{ij}

$|a_{ij}|$ или $\det\|a_{ij}\|$ — определитель квадратной матрицы $\|a_{ij}\|$

$\mathbf{N}$ — множество всех натуральных чисел

$\mathbf{Z}$ — множество всех целых чисел

$\mathbf{Q}$ — множество всех рациональных чисел

$\mathbf{R}$ — множество всех действительных чисел

$\{m, n\}$ или НОК — наименьшее общее кратное чисел m и n

(m, n) или НОД — наибольший общий делитель чисел m и n

$>$, $<$, $\geqslant$, $\leqslant$ — символы «больше», «меньше», «больше или равно», «меньше или равно»

$|a|$ — абсолютное значение (модуль) действительного числа

$\sqrt[n]{\ }$ — корень n-й степени

$\sqrt{\ }$ — квадратный корень

$\log_a M$ — логарифм числа M по основанию a

$\lg M$ — логарифм числа M по основанию 10

$\ln M$ — логарифм числа M по основанию e

$N, n_1 n_2 \ldots n_k \ldots$ — десятичная дробь с целой частью N

$N, n_1 n_2 \ldots n_k (n_{k+1} \ldots n_{k+p})$ — десятичная дробь с периодом $n_{k+1} \ldots n_{k+p}$

z и $\bar{z}$ — комплексно сопряженные числа

i — мнимая единица

$\operatorname{Re} z$ — действительная часть комплексного числа z

$\operatorname{Im} z$ — мнимая часть комплексного числа z

$|z|$ — модуль комплексного числа z

$\operatorname{Arg} z$ — аргумент комплексного числа z

$\boldsymbol{i}$, $\boldsymbol{j}$ — базисные векторы прямоугольной декартовой системы координат на плоскости

$\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ — базисные векторы прямоугольной декартовой системы координат в пространстве

$|\boldsymbol{a}|$ — длина вектора $\boldsymbol{a}$

$(\widehat{\boldsymbol{a}, \boldsymbol{b}})$ — угол между векторами $\boldsymbol{a}$ и $\boldsymbol{b}$

$(\boldsymbol{a}\ \boldsymbol{b})$ — скалярное произведение векторов $\boldsymbol{a}$ и $\boldsymbol{b}$

$\angle$ — символ угла

$\widehat{ABC}$ — величина угла $\angle ABC$

$\parallel$ — символ параллельности

$\perp$ — символ перпендикулярности

$\sphericalangle (a, \alpha)$ — угол между прямой a и плоскостью α

$\angle (\alpha, \beta)$ — угол между плоскостями α и β

$\overset{\smile}{ACB}$ — дуга с концами A и B

$\sin \alpha$ — синус α

$\cos \alpha$ — косинус α

$\operatorname{tg} \alpha$ — тангенс α

$\operatorname{ctg} \alpha$ — котангенс α

$\sec \alpha$ — секанс α

$\operatorname{cosec} \alpha$ — косеканс α

$\arcsin a$ — арксинус числа a

$\arccos a$ — арккосинус числа a

$\operatorname{arctg} a$ — арктангенс числа a

$\operatorname{arcctg} a$ — арккотангенс числа a

(a_n) — последовательность с членами $a_1, a_2, \ldots, a_n, \ldots$

$\lim$ — символ предела

Δx — приращение аргумента

Δf — приращение функции

$f'(x)$ или $\dfrac{df}{dx}$ — первая производная функции $f(x)$

$f^{(n)}(x)$ или $\dfrac{d^n f}{dx^n}$ — n-я производная функции $f(x)$

d — символ дифференциала

$\int$ — символ неопределенного интеграла

$\int_a^b$ — символ определенного интеграла

$\max$ — максимум

$\min$ — минимум

Число $e = 2{,}718281828\ldots$

Число $\pi = 3{,}141592653\ldots$ (π — отношение длины окружности к диаметру)

Число M (модуль перехода) $= 0{,}434294481\ldots$

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Абсолютное значение (модуль) действительного числа 56
— — целого числа 30
Абсцисса точки 256, 259
Аксиома Евклида 222
— линейной полноты 223
Аксиомы арифметики 25
— геометрии 220
— конгруэнтности 221
— непрерывности 223
— порядка 221
— принадлежности 220
Алгебраическая дробь неправильная 94
— — правильная 94
— — простейшая 94
— — рациональная 91, 92
Алгебраическое выражение 80
— — иррациональное 96
— — рациональное 91
— уравнение 99
— число 62
Алгоритм Евклида нахождения НОД двух многочленов 85
— — — — — — чисел 28
Амплитуда колебания 378
Апофема правильного многоугольника 181
— правильной пирамиды 202
— — усеченной пирамиды 203
Аппликата точки 259
Аргумент комплексного числа 77
Арифметическая прогрессия 300
Арифметические неравенства 137
Арифметический корень 57
Арифметическое действие над конечными десятичными дробями 43
Арккосинус 233
Арккотангенс 234
Арксинус 232
Арктангенс 233
Асимптоты 346
Базисные векторы в пространстве 259
— — на плоскости 255
Бесконечная десятичная дробь 42
— числовая последовательность 295
Бесконечно большая величина 320
— малая величина 319
Бесконечное множество 16
— произведение 299
Биекция 15
Биквадратное уравнение 103
Бинарная операция 19
Бином Ньютона 391
Биномиальные коэффициенты 392
Биссектриса треугольника 170
— угла 159
Боковая грань призмы 199
— поверхность конуса 207
— — призмы 199
— — цилиндра 206
Боковое ребро пирамиды 202
— — призмы 200
Боковые стороны трапеции 174

Вектор 269
— нулевой 270
— свободный 270
— связанный 270
— скользящий 270
Векторное произведение 276
Векторы коллинеарные 272
— компланарные 273
Верная цифра приближенного значения числа 65
Вероятность 397
— произведения событий 399
— суммы событий 399
Верхняя подходящая десятичная дробь 55
Вершина многогранника 198
— многогранного угла 197
— многогранной поверхности 198

Вершина параболы 266
— пирамиды 202
— угла 159
Вершины гиперболы 265
— многоугольника 167
— призмы 200
— эллипса 263
Взаимно обратные функции 309
— простые многочлены 82
— — числа 28
Вневписанная окружность 182
Внешние накрест лежащие углы 161
Внешний угол многоугольников 167
Внешняя касательная двух окружностей 178
— область угла 159
Внутренние накрест лежащие углы 161
— точки отрезка 159
Внутренний угол многоугольника 167
Внутренняя касательная двух окружностей 178
— область многогранного угла 198
— — многоугольника 167
— — угла 159
Вогнутость кривой 345
Возвратное уравнение 104
Возрастающая последовательность 282
— функции 307
Вписанный многоугольник 181
— треугольник 182
— угол 179
— четырехугольник 183
Выделение целой части алгебраической дроби 94
Выпуклый многогранник 199
— многогранный угол 198
— многоугольник 167
Высота конуса 207
— пирамиды 202
— призмы 200
— трапеции 174
— треугольника 170
— усеченной пирамиды 203
— цилиндра 206
— шарового сегмента 211
— — сектора 211
— — слоя 212
Вычисление производных элементарных функций 331
Вычитаемое 24, 32
Вычитание дробей 36
— чисел комплексных 72
— — натуральных 24
— — рациональных 38
— — целых 31

Гармонический ряд 296
Геометрическая прогрессия 301
Геометрическое изображение комплексных чисел 75
— множество решений неравенства с двумя неизвестными 147
Гипербола 264
—, директриса 286
—, каноническое уравнение 265
—, фокус 264
—, эксцентриситет 265
Гипотенуза 172
Гомотетия плоскости 218
— пространства 219
Горнера схема 84
Градус 160
Граница области 326
Грань двугранного угла 166
— многогранника 199
— многогранного угла 197
— многогранной поверхности 198
— пирамиды 202
— призмы 199
График функции 303
Группа 20
— аддитивная 20
— коммутативная 20
— мультипликативная 20

Двугранный угол 166
Двучленное уравнение 105
Действительная часть комплексного числа 71
Действительные числа 52
Деление десятичных дробей 44
— многочленов 83
— рациональных дробей 37
— чисел комплексных 72
— — натуральных 24
— — рациональных 38
— — целых 32
Делимое 24
Делитель 24
Десятичная дробь 41
— — бесконечная 42
— — допустимая 42

Десятичная дробь конечная 42
— — периодическая 42
— — подходящая 54
— система счисления 40
Десятичное представление действительного числа 54
Детерминант 128
Диагональ многогранника 199
— многоугольника 168
— параллелепипеда 201
Диагональные матрицы 127
— элементы матрицы 126
Диаметр окружности 176
— сферы 209
Директриса параболы 266
Директрисы гиперболы 266
— эллипса 263
Дискриминант квадратного трехчлена 100
Дифференциал 339
—, геометрическая интерпретация 340
—, инвариантность 341
Дифференциальное уравнение Бернулли 376
— — гармонических колебаний 377
— — линейное 375
— — — с постоянными коэффициентами 379
— — — — — — однородное 379
— — — — — —, частное решение 381
— —, метод вариации постоянных 383
— — однородное 375
— — первого порядка 371
— — с разделяющимися переменными 373
Длина дуги окружности 180
— окружности 180
Додекаэдр 205
Дополнение множества 12
Дополнительные углы 161
Дополнительный множитель 496
Допустимые десятичные дроби 42
— значения переменной 91
Достаточное условие монотонности функции 308
— — сходимости последовательности 288
Достоверное событие 395
Дроби десятичные 41
— неправильные 37
Дроби непрерывные 481
— несократимые 38
— отрицательные 37
— положительные 37
— правильные 37
— рациональные 35
— смешанные 37
Дробно-линейная функция 408
Дробно-линейное неравенство 142
Дуга окружности 178

Единичная окружность 225
Единичное отображение 15
Единичный вектор (орт) 254
— отрезок 254

Зависимая переменная 302
Замкнутая ломаная 166
— область 326
Звенья ломаной 166
Знаки неравенства 137
Знаменатель геометрической прогрессии 301
— дроби 35
Значащая цифра приближенного значения числа 65
Значащие цифры десятичной дроби 42
Золотое сечение 385

Извлечение квадратного корня из действительного числа 68
Измерения параллелепипеда 201
Изометрия 214
Изоморфные множества 19
Икосаэдр 205
Интеграл неопределенный 352
— определенный 364
— с переменным верхним пределом 366
Интегральная кривая 372
— сумма 364
Интегрирование, метод замены переменной 356
—,— интегрирования по частям 353
—, подстановки Эйлера 359
Инъективное отображение 14
Иррациональное выражение 96
— неравенство 143
— уравнение 107
— число 51, 59

Исключение иррациональности из числителя (знаменателя) дроби 97
Испытание 395
Исход 395

Каноническое разложение натуральных чисел 26
— уравнение гиперболы 265
— — параболы 266
— — прямой 268
— — эллипса 263
Касательная к окружности 177
Катет 172
Квадрант 227
Квадрат 174
Квадратичная функция 409
Квадратная матрица 124
Квадратное неравенство 141
— уравнение 100
Квадратный многочлен 80
Классификация разрывов функции 324
Коллинеарные векторы 272
Кольцо 21
Компланарные векторы 273
Комплексно сопряженные числа 73
Комплексное число 71
— — , аргумент 77
— — , возведение в степень 78
— — , геометрическое изображение 75
— —, действительная и мнимая части 71
— —, извлечение корня 78
— —, модуль 73
— —, тригонометрическая форма 77
Композиция преобразований 214
Конечная непрерывная дробь 50
Конечное множество 16
Конус 207
— усеченный 208
Концы отрезка 159
Координатная ось 254
— плоскость 255
Координаты вектора в пространстве 275
— — на плоскости 274
— точки в пространстве 259
— — на плоскости 256
Корень из действительного числа 57
Корень многочлена 86
— уравнения 98
Косеканс 227
Косинус 226
Кососимметрическая матрица 127
Кососимметрическое уравнение 104
Котангенс 226
Коэффициент гомотетии 218
— подобия 219
Коэффициенты уравнения 98, 99
Кратный корень многочлена 86
Криволинейная трапеция 363
Критерий Коши 317
Критические точки функции 342
Круг 177
Куб 201
Кубический многочлен 80

Лексикографический способ записи многочлена 90
Линейная система уравнений 117
— функции 406
Линейное уравнение 100
Линейный многочлен 80
— угол двугранного угла 166
Линия второго порядка 259
— первого порядка 259
— центров 178
Логарифм 58
— десятичный 58
— натуральный 58
Логарифмическая функция 412
Логарифмическое неравенство 145
— уравнение 114
Ломаная 166
— замкнутая 166
— простая 166
Луч 174

Максимум функции 342
Математической индукции метод 384
— — принцип 384
Матрица 124
Мгновенная скорость 330
Мгновенное ускорение 335
Медиана 169
Метод Гаусса 117
— дополнительного угла 245
— замена переменной 356

Метод интегрирования по частям 353
— интервалов 142
— касательных (метод Ньютона) 66
— математической индукции 384
Минимум функции 342
Минор 144
Минута 160
Мнимая единица 71
— часть комплексного числа 71
Многогранная поверхность 198
Многогранник 198
— правильный 203
Многогранный угол 197
Многоугольник 167
—, вписанный в окружность 181
— выпуклый 167
—, описанный около окружности 181
— правильный 168
Многочлен (полином) 80, 89
— квадратный 80
— кубичный 80
— линейный 80
— неприводимый 89
— однородный 90
—, разложение на множители 88
Многочлены взаимно простые 82
Множества бесконечные 16
— конечные 16
— несчетные 16
— равные 11
— решений неравенства 140
— счетные 16
— упорядоченные 13
— эквивалентные 15
Множество 10
— допустимых значений неизвестной 99
— значений функции 302
— пустое 10
Множитель дополнительный 106
Модуль вектора 269
— действительного числа 56
— комплексного числа 73
Монотонные последовательности 282
— функции 307

Наибольшее значение функции 344
Наибольший общий делитель многочленов 82
— — — чисел 28
Наименьшее значение функции 344
— общее кратное 27
Наклонная 164
Направление вектора 269
— в пространстве 158
— на плоскости 158
Направляющие косинусы 275, 276
Натуральные числа 23
Натуральный логарифм 58
Начало координат 254
Начальная фаза колебания 378
Невозможные события 395
Невозрастающая последовательность 282
— функция 307
Независимая переменная 302
Независимые события 395
Неизвестное 98
Необходимое и достаточное условие существования предела функции 317
— — — — сходимости положительного ряда 295
— условие сходимости бесконечного числового ряда 295
— — — последовательности 292
Неоднородная система уравнений 117
Неопределенная система уравнений 117
Неопределенный интеграл 352
Неправильная алгебраическая дробь 94
— рациональная дробь 37
Непрерывная дробь 48
Непрерывность функции 322, 326
Неравенства 137
— арифметические 137
— дробно-линейные 142
— иррациональные 143
— квадратные 141
— линейные 140
— логарифмические 145
— показательные 144
— тригонометрические 246
Неравенство Бернулли 384
— Гёльдера 139
— Коши 139
— Коши — Буняковского 139

Неравенство треугольника 168
Несовместные события 395
Несократимая дробь 38
Нестрогие неравенства 137
Несчетное множество 17
Неубывающая последовательность 282
— функция 307
Нечетная перестановка 386
— подстановка 389
— функция 305
Нечетное число 27
Нижняя подходящая десятичная дробь 54
Нормальное уравнение плоскости 267
Нулевой вектор 270

Область изменения функции 302
— определения функции 302, 405
Образ элемента 14
Образующая конуса 207
— усеченного конуса 209
— цилиндра 206
Обратная подстановка 389
— функция 309
Обратно пропорциональная зависимость 407
Обратное отображение 15, 213
— число 53
Обратные тригонометрические функции 232
Обращение десятичной дроби в рациональную дробь 44
Общее кратное чисел 27
Общий делитель чисел 28
— перпендикуляр двух скрещивающихся прямых 164
Объединение множеств 11
— событий 396
Объем додекаэдра 205
— икосаэдра 205
— конуса 208
— куба 201
— октаэдра 204
— параллелепипеда 201
— пирамиды 202
— призмы 200
— тел вращения 369
— тетраэдра 204
— усеченного конуса 209
— усеченной пирамиды 203
— цилиндра 207
— шара 210
Объем шарового сегмента 211
— — сектора 211
— — слоя 212
Обыкновенное дифференциальное уравнение 370
Ограниченная функция 307
Однородная система уравнений 117
Однородный многочлен 90
Односторонний предел функции 314
Односторонняя непрерывность функции 324
Одночлен 80, 89
Округление числа с избытком 66
— — — недостатком 66
Окружность 176
— вневписанная 182
— , вписанная в многоугольник 181
—, — — треугольник 182
—, описанная около многоугольника 181
—, — — треугольника 182
Октаэдр 204
Описанный многоугольник 181
— треугольник 182
— угол 180
— четырехугольник 183
Определенный интеграл 364
Определитель 128
Ордината точки 256, 259
Ортогональная проекция 165
Ортоцентр треугольника 170
Осевая симметрия плоскости 216
— — пространства 216
Осевое сечение 217
Основание конуса 207
— логарифма 58
— пирамиды 202
— призмы 199
— степени 25, 57
— трапеции 174
— треугольника 170
— цилиндра 206
— шарового сегмента 211
— — слоя 212
Основная матрица 132
— теорема алгебры 87
— — арифметики 26
— формула интегрального исчисления 367
Основные свойства неравенства 138

Основные свойства производных 330
Остаток от деления многочленов 83
— ряда 294
Остроугольный треугольник 169
Острый угол 160
Ось абсцисс 256
— аппликат 259
— вращения 217
— координатная 254
— ординат 256
— симметрии 216
Открытая область 326
Открытый луч 157
Относительная погрешность 64
Отношение крайнего и среднего 385
— порядка в множестве действительных чисел 52
— эквивалентности (множеств) 15
Отображение взаимно однозначное 15
— множества в множество 14
— — на множество 14
— обратное 15
Отрезок 159
— натурального ряда 16
Отрицательная полуось 56
Отрицательное направление вращения 225
Отрицательные углы 225
— целые числа 30

Парабола 266, 409
Параллелепипед 200
— прямой 201
— прямоугольный 201
Параллелограмм 172
Параллельная проекция 165
Параллельные плоскости 163
— прямые 161
Параллельный перенос 270
Параметрическое уравнение прямой 268
Первообразная 352
Переменная зависимая 302
— независимая 302
Перемещение (изометрия) 213
Пересечение множеств 11
— событий 397
Перестановка 386
— с повторениями 392
Период 324
— десятичной дроби 42
Периодическая функция 306
Перпендикулярность прямой и плоскости 163
Перпендикулярные плоскости 164
— прямые 162
Пирамида 201
— правильная 202
— усеченная 203
Плоский угол многогранного угла 198
Плоскостные аксиомы равенства 222
Плоскость симметрии 218
Площадь боковой поверхности конуса 208
— — — пирамиды 202
— — — призмы 200
— — — цилиндра 207
— квадрата 174
— криволинейной трапеции 363
— круга 180
— параллелограмма 173
— поверхности додекаэдра 205
— — икосаэдра 204
— — октаэдра 204
— — тетраэдра 204
— — усеченного конуса 209
— — усеченной пирамиды 203
— — шарового пояса 212
— — — сектора 211
— правильного многоугольника 181
— прямоугольника 174
— прямоугольного треугольника 172
— ромба 173
— сегмента 180
— сегментной поверхности 211
— сектора 181
— сферы 210
— трапеции 175
— треугольника 169
— шарового пояса 212
Поверхность многогранника 199
Поворот плоскости вокруг точки 215
Погрешность абсолютная 63
— — предельная 64
— относительная 64
— — предельная 64
Подмножество 10
— собственное 11

Подобие 219
— фигур 220
Подобные многоугольники 175
— одночлены 80
Подстановки 388
— Эйлера 359
Подходящая дробь 50
Подынтегральная функция 352, 364
Подынтегральное выражение 352, 364
Показатель степени 25
Показательная функция 411
Показательное неравенство 144
— уравнение 113
Поле 21
Полиномиальные коэффициенты 394
Полная группа событий 396
— — — несовместных 396
Положительная полуось 56
Положительное направление вращения 225
Положительные углы 225
— целые числа 30
Положительный числовой ряд 295
Полюс 257
Полярная ось 257
— система координат 257
Полярные координаты 257
Полярный радиус 257
— угол 257
Порядок дифференциального уравнения 370
Последовательность бесконечная 276
— возрастающая 282
— конечная 279
— монотонная 282
— ограниченная 281
— расходящаяся 284
— сходящаяся 284
— убывающая 282
— Фибоначчи 385
Построение графика функции 347
— касательных к окружности 191
— окружностей и дуг окружностей 189
— окружности, вписанной в многоугольник, и многоугольника описанного около окружности 193
Построение окружности, описанной около многоугольника 192
— отрезков 186
— прямых, параллельных и перпендикулярных данной прямой 184
— треугольников 194
— углов 185
Правило параллелепипеда 273
— параллелограмма 271
— треугольника 272
Правильная алгебраическая дробь 94
— призма 200
— рациональная дробь 37
Правильный многогранник 203
— многоугольник 168
Предел последовательности 284
— функции 312
— — двух переменных 326
Пределы интегрирования 364
Преобразование графика функции 404
— плоскости (пространство) 213
— подобия 213
Приближенное значение числа 63
Приведенное квадратное уравнение 100
Призма 199
— правильная 200
Признак параллельности плоскостей 163
— — прямой и плоскости 163
— — прямых 161
— перпендикулярности плоскостей 164
— — прямой и плоскости 163
— сходимости числовых рядов Даламбера 297
— — — — Коши 297
— — — — Раабе 297
Признаки делимости чисел 27
— подобия многоугольников 175
— — треугольников 175
Прилежащие углы 161
Принцип вложенных отрезков 54
— математической индукции 384
Приращение аргумента функции 327
— функции 327
Прогрессия арифметическая 300
— геометрическая 301

Проекция ортогональная 165
— параллельная 165
Произведение бесконечное 299
— вектора на число 272
— матриц 126
— многочленов 81
— подстановок 389
— последовательностей 280
— рациональных дробей 36
— событий 397
— функций 304
— чисел комплексных 72
— — натуральных 23
— — рациональных 38
— — целых 31
Производная 328
—, геометрическая интерпретация 329
—, кинематический смысл 330
Производные высшего порядка 334
— обратных тригонометрических функций 333
— показательной и логарифмической функций 331
— пропорции 59
— степенной функции 332
— тригонометрических функций 333
— функции, заданной в параметрической форме 336
Промежуток монотонности функции 308
Прообраз 14
— полный 14
Пропорциональные отрезки 161
Пропорция 58
Простая ломаная 166
Простейшая алгебраическая дробь 94
Простейшие правила интегрирования 353
Простое число 26
Простой корень многочлена 86
Противоположно направленные лучи 158
Противоположное комплексное число 72
Противоположные векторы 269
— события 395
Прямая призма 200
— пропорциональная зависимость 407
Прямой круговой конус 207
— — цилиндр 206
Прямой параллелепипед 201
— угол 160
Прямоугольная декартова система координат в пространстве 258
— — — — на плоскости 255
Прямоугольник 173
Прямоугольный параллелепипед 201
— треугольник 172

Равенство векторов 270
Равнобедренная трапеция 174
Равнобедренный треугольник 171
Равновеликие треугольники 169
Равновероятные события 398
Равносторонний треугольник 171
Равные углы 159
Радиан 160
Радиус вневписанной окружности 183
— окружности 176
— —, вписанной в треугольник 182
— —, описанной около треугольника 182
— основания конуса 207
— — цилиндра 206
— шара 209
Развернутый угол 159
Разложение вектора 273
— многочлена на множители 88
— чисел на множители 26
— — — простые множители 26
Размещения 390
Разность арифметической прогрессии 300
— векторов 271
— дробей 36
— матриц 125
— многочленов 82
— множеств 12
— последовательностей 280
— событий 397
— функций 304
— чисел комплексных 72
— — натуральных 24
— — рациональных 38
— — целых 31
Ранг матрицы 130
Расстояние между скрещивающимися прямыми 164

Расстояние от точки до плоскости 164, 268
— — — — прямой 162, 261
Расходящаяся последовательность 284
Расширенная матрица 132
Рациональная алгебраическая дробь 91
— дробь 35
— степень действительного числа 57
Рациональное алгебраическое уравнение 99
— число 38
Ребро двугранного угла 166
— многогранного угла 197
— многогранной поверхности 198
— пирамиды 202
— призмы 200
Решение дифференциального уравнения 371, 377
— системы уравнений 116
Решения уравнения 98
Ромб 173
Ряд гармонический 296
— положительный 295
— числовой 293

Свободный вектор 270
Свойство Архимеда 53
— непрерывности множества действительных чисел 53
— ортогональной проекции плоского многоугольника 165
Сегмент 197
Секанс 227
Сектор 181
Секунда 160
Секущая 177
Симметрическая матрица 127
Симметрическое уравнение 104
Симметрия осевая 216
— относительно плоскости 216
— центральная 215
Синус 226
Система вложенных отрезков 53
— координат полярная 257
— — прямоугольная в пространстве 258
— — — на плоскости 255
— счисления 40, 412
— — восьмеричная 413
— — двоичная 415
— — десятичная 40
Система уравнений 116
— — линейных 117
— — — неоднородная 117
— — — неопределенная 117
— — — несовместная 116
— — — однородная 117
— — — совместная 117
Скалярное произведение векторов 274
Скалярный квадрат (вектора) 274
Скрещивающиеся прямые 164
Сложение векторов 271
— вероятностей 398
— многочленов 81
— рациональных дробей 36
— чисел комплексных 72
— — натуральных 23
— — рациональных 38
— — целых 31
Сложная функция 305
Случайное событие 395
Смежные звенья ломаной 166
— углы 161
Смешанная дробь 37
— — периодическая десятичная 42
Смешанное произведение векторов 277
Собственный максимум 350
— минимум 350
Сокращение дробей 35
Сонаправленные лучи 158
Соответственные углы 161
Соответствие 14, 302
Составное число 26
Сочетания 390
— с повторениями 394
Среднее арифметическое 139
— гармоническое 296
— геометрическое 139
Средние члены пропорции 58
Средняя линия трапеции 191
— — треугольника 171
Стационарная точка функции 344
Степенная функция 411
Степень действительного числа 57
— комплексного числа 78
— многочлена 80, 89
— одночлена 80, 89
— уравнения 98
Стерадиан 212
Стороны угла 159

Строгие неравенства 137
Сумма векторов 271
— матриц 125
— многочленов 81
— последовательностей 280
— рациональных дробей 36
— событий 425
— углов 159
— функций 304
— чисел комплексных 72
— — натуральных 23
— — рациональных 40, 41
— — целых 31
— числового ряда 294
— членов арифметической прогрессии 300
— — бесконечно убывающей геометрической прогрессии 301
— — геометрической прогрессии 301
Сфера 209
Схема Горнера 84
Сходящаяся последовательность 284
Сходящийся ряд 294

Тангенс 226
Телесный угол 212
Теорема Безу 85
— Больцано — Коши 293
— Вейерштрасса 288
— косинусов 250
— Кронекера — Капелли 132
— Лагранжа 331
— Пифагора 172
— Ролля 331
— синусов 250
— Фалеса 161
— Ферма 331
Теоремы о параллельных плоскостях 163
— — перпендикулярности прямой и плоскости 163
— — плоскости и прямой, параллельной плоскости 163
— — пределах последовательностей 303
— — — функций 315
— — производных 347
— — пропорциональных отрезках 162
— — равных отрезках 161
Тетраэдр 204
Тождественная подстановка 389
Тождественное отображение 15
Тождественное преобразование 214
Тождество 98
Точка перегиба 345
— приложения вектора 270
Транспозиция перестановок 387
Трансцендальная форма системы линейных уравнений 120
Трансцендентное уравнение 113
— число 62
Трапеция 174
— криволинейная 363
— равнобочная 174
Треугольная форма системы линейных уравнений 120
Треугольник 168
— , вписанный в окружность 182
—, описанный около окружности 182
— остроугольный 169
— Паскаля 391
— прямоугольный 169
— равнобедренный 171
— равносторонний 171
— тупоугольный 169
Тригонометрическая форма записи комплексного числа 77
— функция 225
Тупой угол 160

Убывающая последовательность 282
Убывающие функции 308
Угловой коэффициент прямой 260
Углы внешние накрест лежащие 161
— внутренние накрест лежащие 161
— дополнительные 161
— прилежащие 161
— равные 159
— смежные 161
— соответственные 161
Угол 159
— , вписанный в окружность 179
— двугранный 166
— линейный двугранного угла 166
— между векторами 274
— — направлениями 161
— — плоскостями 166
— — прямой и плоскостью 165
— многогранный 197

Угол, описанный около окружности 180
— острый 160
— отрицательный 225
— поворота 215
— положительный 225
— прямой 160
— тупой 160
— центральный 178
Уменьшаемое 24
Умножение вектора на число 272
— десятичных дробей 43
— матриц 126
— матрицы на число 125
— многочленов 81
— подстановок 389
— рациональных дробей 36
— чисел комплексных 72
— — натуральных 23
— — рациональных 38
— — целых 31
Упорядоченное множество 13
Уравнение 98
— алгебраическое 99
— биквадратное 103
— возвратное 104
— гармонических колебаний 377
— двучленное 105
— диофантово 112
— дифференциальное 370
— иррациональное 107
— касательной 330
— квадратное 100
— кососимметрическое 104
— кубическое 102
— линейное 100
— логарифмическое 114
— окружности 262
— плоскости 267
— показательное 113
— прямой 260
— рациональное 107
— симметрическое 104
— трансцендентное 113
— тригонометрическое 241
Уравнения эквивалентные 93
Усеченная пирамида 203
Усеченный конус 209
Условие монотонности функции 341
— параллельности прямой и плоскости 269
— — прямых 262, 269
Условие перпендикулярности прямой и плоскости 269
— — прямых 262, 269
Условная вероятность 400

Факториал 387
Фигура 213
— вращения 217
Фигуры гомотетичные 219
— подобные 220
Фокальные радиусы гиперболы 264
— — эллипса 263
Фокус параболы 266
Фокусы гиперболы 264
— эллипса 262
Формула Бернулли 402
— Муавра 78
— общего члена арифметической прогрессии 300
— — — геометрической прогрессии 301
— — — последовательности 280
— полной вероятности 401
— Эйлера 205
Формулы Виета 87
— Крамера 132
— приведения 237
— сокращенного умножения 86
Функция монотонная 307
— нескольких независимых переменных 310
— нечетная 305
— обратная 309
— ограниченная 307
— периодическая 306
— четная 305
— числовая 302

Характеристическое уравнение 380
Хорда окружности 176
— сферы 209

Целая часть алгебраической дроби 94
— — десятичной дроби 41
— — рациональной дроби 37
Центр гиперболы 265
— гомотетии 234
— окружности 176
— правильного многоугольника 181
— — треугольника 199
— симметрии 231

Центр шара 209
— эллипса 263
Центральная симметрия 215
Цилиндр 205
Цилиндрическая поверхность 205
Цифры 40
— верные 65
— значащие 65

Частичная сумма ряда 293
Частичное произведение 299
Частная производная 337
Частное многочленов 83
— последовательностей 280
— приращение 337
— рациональных дробей 37
— функций 304
— чисел комплексных 72
— — натуральных 24
— — рациональных 38
— — целых 32
Частота событий 397
Четверти 227
Четная перестановка 387
— подстановка 389
— функция 305
Четное число 27
Четырехугольники, вписанные в окружность 183
—, описанные около окружности 183
Числа алгебраические 62
— взаимно простые 28
— действительные 52
— иррациональные 51, 59
— комплексные 71
— мнимые 71
Числа натуральные 23
— нечетные 27
— отрицательные 30
— положительные 30
— простые 26
— противоположные 30
— рациональные 38
— составные 26
— трансцендентные 62
— целые 30, 32
— четные 27
Числитель дроби 35
Число e 291
Числовая последовательность. 279
Числовой ряд 293
Член последовательности 279
— ряда 293

Шар 209
Шаровой пояс 212
— сегмент 211
— сектор 211
— слой 211

Эквивалентные дроби 35
— множества 15
— неравенства 140
— уравнения 99
Экстремум функции 342
— — двух переменных 350
Эксцентриситет гиперболы 265
— эллипса 263
Элемент матрицы 124
— множества 10
Элементарные (простейшие) дроби 94
— события 396

Александр Геннадьевич Цыпкин
СПРАВОЧНИК ПО МАТЕМАТИКЕ
для средних учебных заведений

Редакторы **Е. И. Стечкина, Т. В. Шароватова**
Художественный редактор **Г. М. Коровина**
Технический редактор **И. Ш. Аксельрод**
Корректоры **Е. Ю. Рычагова, Н. Б. Румянцева**

ИБ № 32296

Сдано в набор 28.01.87. Подписано к печати 31.08.87. Формат 84 x 108/32. Бумага тип. № 3. Литературная гарнитура. Высокая печать. Усл. печ. л. 22.68. Усл. кр.-отт. 22,89. Уч.-изд. л. 30.18. Тираж 300 000 экз. (2-й завод 100001—300000 экз.).
Заказ № 8016. Цена 1 р. 70 к.

Ордена Трудового Красного Знамени
издательство «Наука»
Главная редакция
физико-математической литературы
117071, Москва В-71, Ленинский проспект, 15

Ордена Октябрьской Революции и ордена Трудового Красного Знамени МПО «Первая Образцовая типография» имени А. А. Жданова. Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли.
113054, Москва В-54, Валовая, 28.

Отпечатано в типографии «Коммуна»
г. Воронеж, пр. Революции, 39.